

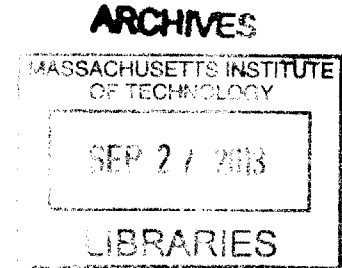
Justifying Bayesianism

by

Jennifer Rose Carr

B.A., Stanford University, 2006

M.A., Harvard University, 2008



Submitted to the Department of Linguistics and Philosophy
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2013

© Massachusetts Institute of Technology 2013. All rights reserved.

Author _____

Department of Linguistics and Philosophy
August 12, 2013

Certified by _____

Richard Holton
Professor of Philosophy
Dissertation Supervisor

Accepted by _____

Alex Byrne
Professor of Philosophy
Chair, Departmental Committee of Graduate Studies

Justifying Bayesianism

Jennifer Carr

Submitted to the Department of Linguistics and Philosophy on September 4, 2013
in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Abstract

Bayesianism, in its traditional form, consists of two claims about rational credences. According to the first claim, *probabilism*, rational credences form a probability function. According to the second claim, *conditionalization*, rational credences update by conditionalizing on new evidence. The simplicity and elegance of classical Bayesianism make it an attractive view. But many have argued that this simplicity comes at a cost: that it requires too many idealizations.

This thesis aims to provide a justification of classical Bayesianism. Chapter One defends probabilism, classically understood, against the charge that by requiring credences to be precise real numbers, classical Bayesianism is committed to an overly precise conception of evidence. Chapter Two defends conditionalization, classically understood, against the charge that epistemic rationality consists only of synchronic norms. Chapter Three defends both probabilism and conditionalization against the objection that they require us, in some circumstances, to have credences that we can know are not as close to the truth as alternatives that violate Bayesian norms.

Thesis Supervisors: Richard Holton

Title: Professor of Philosophy

Acknowledgments

For valuable feedback, thanks to Selim Berker, Rachael Briggs, Alan Hájek, Agustín Rayo, Robert Stalnaker, and audiences at the Australian National University, Washington University, University of Missouri, Columbia, University of Leeds, University of Bristol, and MIT.

Profound gratitude to Selim Berker and Susanna Siegel for inviting an interested literature grad student into philosophy. The two of you changed my life dramatically; thanks for welcoming me into the fold.

I cannot thank my committee warmly enough: many thanks to Steve Yablo, for three-hour meetings and questions that always manage to stump me; to Roger White, for devoting so much time and attention to a dissertation that veered into epistemology late in its life (and for so many hilarious counterexamples); and to Richard Holton, a wonderful, wonderful advisor, without whose gentle prodding and weekly feedback I never would have given so many talks, never would have finished my degree in five years, and never would have had such a lively and happy time writing a dissertation at MIT.

For years of discussion, suggestions, objections, and beer-swilling, thanks to the entire MIT graduate student community, and especially Ryan Doody, David Gray, Daniel Greco, Daniel Hagen, Sophie Horowitz, Sofia Ortiz-Hinojosa, Damien Rochford, Bernhard Salow, and Paolo Santorio.

And finally, thanks with love to my family.

Contents

1	Imprecise Evidence without Imprecise Credences	11
1	The imprecise view	12
1.1	Some examples	12
1.2	What are imprecise credences?	15
2	Challenges for imprecise credences	16
2.1	Fending off the permissivist alternative	16
2.2	The pragmatic challenge	17
2.3	The epistemic challenge	19
3	A poisoned pawn	21
4	The precise alternative	24
5	The showdown	27
5.1	Nonsharper claim #1: Precise credences should reflect known chances	28
5.2	Nonsharper claim #2: Precise credences are “too informative”	30
5.3	Nonsharper claim #3: Imprecise confirmation requires imprecise credences	33
6	Conclusion	37
2	Don’t Stop Believing	39
1	The conflict	41
1.1	Bayesianism	41
1.2	The rejection of diachronic rationality	42
1.3	Orienting the debate	45
2	Problems for time-slice rationality	46
2.1	Problem #1: permissibly discarding evidence	46
2.2	Problem #2: deviating from epistemic ideals	47
2.3	Problem #3: all incoherence is diachronic	49
3	Epistemic ‘blamelessness’ does not entail epistemic ideality	50

3.1	Diachronic evidentialism and information loss	50
3.2	Epistemic <i>ought</i> implies <i>can</i> ?	52
3.3	Relative rationality	53
3.4	Rationality vs. epistemic ideality?	54
4	Rational information loss	55
4.1	Losing information to gain information	55
4.2	Epistemic utility theory	57
4.3	Assessing rational trade-offs	58
4.4	Discussion	60
5	Conclusion	61
3	What to Expect When You're Expecting	64
1	Epistemic utility theory	66
2	When the world isn't independent of our beliefs	68
3	Conditionalization and expected accuracy	71
3.1	Two kinds of decision theory	71
3.2	Causal expected accuracy	72
3.3	Evidential expected accuracy	74
3.4	Discussion	75
3.5	Expected accuracy in epistemic utility theory	76
4	Probabilism and expected accuracy	78
5	What is observational expected accuracy?	80
5.1	OEA's idealizing assumption	80
5.2	The standard motivations for MaxExAcc don't motivate maximizing OEA	81
5.3	The truth independent of our credences?	83
5.4	Expected accuracy of credal acts	85
6	Problems for accuracy dominance arguments	87
6.1	Joyce's "dominance" argument	87
6.2	What this shows about epistemic utility functions	90
7	Accuracy and evidence	92

Introduction

Bayesianism, in its traditional form, consists of two claims about rational credences. According to the first claim, *probabilism*, rational credences have a particular mathematical structure at each point in time: they form a probability function. According to the second claim, *conditionalization*, rational credences are interrelated in a particular way over time: they update by conditionalizing on new evidence.

The simplicity and elegance of classical Bayesianism make it an attractive view. But many have argued that this simplicity comes at a cost: that it requires too much idealization. One objection claims that by requiring credences to be precise real numbers, classical Bayesianism ignores the fact that our evidence is messy and imprecise. Another objection claims that our credences should only be constrained by evidence available at the present, and so seeks to unseat cross-temporal constraints like conditionalization. Finally, a third objection claims that rational credences must approximate the truth, and both probabilism and conditionalization will sometimes require agents to have credences that are predictably farther from the truth than non-probabilistic or non-conditionalized alternatives.

This thesis aims to provide a justification of classical Bayesianism against each of these challenges.

Chapter One offers a defense of probabilism, classically understood. Some philosophers claim that, in response to ambiguous or unspecific evidence, rationality requires adopting imprecise credences: degrees of belief that are spread out over multiple real numbers. Otherwise, on this view, one's credences would take an inappropriately definite stance on the basis of indefinite evidence. I argue that these views conflate two kinds of states: the state of having a precise credence in a proposition, and the state of having a belief about that proposition's objective or epistemic probability. With this distinction in hand, there are a variety of positions open to the defender of precise credences that are all compatible with the idea that evidence can be unspecific or ambiguous.

Chapter Two offers a defense of conditionalization, classically understood. In

recent years it's often been argued that the norms of rationality only apply to agents at a time; there are no diachronic constraints on rationality. Some have claimed that we should replace conditionalization, a diachronic norm, with some similar norm that only requires us to constrain our credences to the evidence that is currently accessible to us. I show that abandoning diachronic norms incurs serious costs for a theory of rationality, and I argue that the motivations for preferring a synchronic-norms-only view rest on a misguided idea that agents are responsible or blameworthy for whether their credences accord with epistemic norms.

Chapter Three considers an objection to both probabilism and conditionalization, which comes by way of a form of epistemic decision theory. On this style of decision theory, rational credences are those that best approximate the truth. Both probabilism and conditionalization require us, in some circumstances, to have credences that we can know are not as close to the truth as alternatives that violate probabilism or conditionalization. I argue in favor of another mathematical apparatus for assessing credences in terms of epistemic utility, one that is importantly different from its decision-theoretic counterpart. But, I argue, this other apparatus is widely misinterpreted as recommending credences that best approximate the truth. It doesn't; but it's not clear what it does do. This other mathematical apparatus therefore stands in need of a philosophical interpretation.

Chapter 1. Imprecise Evidence without Imprecise Credences

Rationality places constraints on our beliefs. We should have the beliefs that our evidence entails; we should have the credences (degrees of belief or confidence) that our evidence supports. And rationality can place requirements on the fineness of grain of our belief states. Proponents of precise credences (“sharppers”) hold that rationality requires agents to have attitudes that are comparatively fine-grained: credences that are each represented by a unique real number. Proponents of imprecise or “mushy” credences (“nonsharppers”) hold that, in response to some kinds of evidence, rationality requires credences that are coarser-grained, spread out over multiple real numbers.¹

Who’s right?

It’s important to distinguish the question of what credences we *should* have from the question of what credences we do have. Even the sharper can agree that precise credences might not be psychologically realistic for a number of reasons. And so it might be that actual agents have, by and large, imprecise credences. But this descriptive observation is orthogonal to the normative question that is here at issue.

This paper concerns, instead, the nature of the norms that rationality imposes on us. Does rationality require imprecise credences? Nonsharppers hold that it does: ambiguous or unspecific evidence requires correspondingly ambiguous or unspecific credences. I will argue that this is false. Ambiguous or unspecific evidence, if it exists, at most requires *uncertainty* about what credences to have. It doesn’t require credences that are themselves ambiguous or unspecific.

Part of what is at stake in answering this question is the viability of the ar-

¹ A third view might say that imprecise credences are sometimes permissible but never required. I’ll ignore this view in my discussion, but for taxonomical purposes, I understand this to be a precise view. In my taxonomy, it’s essential to the nonsharper view that some bodies of evidence *mandate* imprecise credences, rather than simply permitting them.

ray of tools that have been developed within the orthodox probabilistic framework, traditional Bayesianism. Dropping traditional Bayesianism requires starting from scratch in building decision rules (norms about what choices are rational) and update rules (norms about how our beliefs should evolve in response to new information). And as we'll see, proposed replacements for the traditional decision rules and update rules have serious costs, including permitting what is intuitively rationally impermissible, and prohibiting what is intuitively rationally permissible.

In sections 1 and 2, I introduce the imprecise view, its intuitive appeal, and what I take to be its toughest challenges. In section 3, I discuss an attractive strategy for avoiding these challenges. But once the veil is lifted, the strategy is revealed to be a notational variant of a precise view. On this precise view, which I lay out in section 4, instead of representing agents as *having* multiple probability functions, we think of agents as being *uncertain* over multiple probability functions. This precise view can accommodate all of the (good) motivations for the imprecise view but faces none of its challenges.

In section 5 we finally reach the showdown. Are there any reasons to adopt the imprecise view that aren't equally good reasons for adopting the precise alternative I laid out in section 4? The answer, I argue, is *no*: anything mushy can do, sharp can do better.

1 The imprecise view

1.1 Some examples

Traditional Bayesians hold that beliefs come in degrees, conventionally represented as real numbers from 0 to 1, where 1 represents the highest possible degree of confidence and 0 represents the lowest. An agent's degrees of belief, standardly called "credences," are constrained by the laws of probability and evolve over time by updating on new evidence.

Nonsharppers hold that in the face of some bodies of evidence, it is simply irrational to have precise credences. These bodies of evidence are somehow ambiguous (they point in conflicting directions) or unspecific (they don't point in any direction). It's an open question how widespread this kind of evidence is. On some versions of the view, we can only have precise credences if we have knowledge of objective chances. And so any evidence that doesn't entail facts about objective chances is ambiguous or unspecific evidence, demanding imprecise credences.²

² Note: When I speak of having imprecise credences in light of bodies of evidence, I'm including empty or trivial bodies of evidence.

A variety of cases have been presented to elicit the nonsharper intuition. In these cases, the nonsharper says, any precise credence would be unjustified or irrational. Here are a few examples:

Toothpaste/jellyfish

“A stranger approaches you on the street and starts pulling out objects from a bag. The first three objects he pulls out are a regular-sized tube of toothpaste, a live jellyfish, and a travel-sized tube of toothpaste. To what degree should you believe that the next object he pulls out will be another tube of toothpaste?” (Elga, 2010, 1)

If there’s any such thing as unspecific or ambiguous evidence, this looks like a good candidate. Unless you have peculiar background beliefs, the evidence you’ve received can seem too unspecific to support any particular precise credence. It doesn’t obviously seem to favor a credence like .44 over a credence like .21 or .78. So what should you do when you receive evidence like this?

There’s something puzzling about the idea that there could be a unique degree to which your body of evidence confirms the hypothesis that the next object pulled from the bag will be a tube of toothpaste. (What would it be?) So maybe neutrality demands that you take on a state that equally encompasses *all* of the probability functions that could be compatible with the evidence.

Here is a second example that has been used to motivate imprecise credences:

Coin of unknown bias

You have a coin that you know was made at a factory where they can make coins of pretty much any bias. You have no idea whatsoever what bias your coin has. What should your credence be that when you toss it, it’ll land heads? (See e.g. Joyce 2010.)

This sort of case is somewhat more theoretically loaded. After all, there is a sharp credence that stands out as a natural candidate: .5. But Joyce (2010) and others have argued that the reasoning that lands us at this answer is faulty. The reasoning relies on something like the principle of indifference (POI). According to POI, if there is a finite set of mutually exclusive possibilities and you have no reason to believe any one more than any other, then you should distribute your credence equally among them. But POI faces serious (though arguably not decisive) objections.³ Without something like it, what motivates the .5 answer?

³ In particular, the multiple partitions problem; see e.g. van Fraassen’s (1989) cube factory example.

According to Joyce, nothing does. There's no more principled reason to settle on a credence like .5 than a credence like .8 or .4211. Joyce sees this as a case of unspecific evidence. Of course, if you knew the precise objective chance of the coin's landing heads, then you should adopt that precise credence. But if you have no information at all about the objective chance, then the rational thing to do is to have a credence that represents all of the numbers that could, given your evidence, be equal to the objective chance of the coin's landing heads. In this case, that might be the full unit interval $[0, 1]$. On Joyce's view, adopting any precise credence would amount to making some assumptions that are unwarranted by your very sparse evidence. (More on this in section 5.)⁴

The general claim—some might even say *intuition*—that underpins the non-sharper's assessment of these sorts of cases is: any precise credence function would be an inappropriate response to the evidence. It would amount to “taking a definite stance” when the evidence doesn't justify a definite stance. Or it would involve adopting attitudes that are somehow much more informed or informative than what the evidence warrants. Or it would involve failing to fully withhold judgment where judgment should be withheld.

Some quotations from defenders of imprecise credences:

Precise credences... always commit a believer to extremely definite beliefs about repeated events and very specific inductive policies, even when the evidence comes nowhere close to warranting such beliefs and policies. (Joyce, 2010, 285)

If you regard the chance function as indeterminate regarding X , it would be odd, and arguably irrational, for your credence to be any sharper... How would you defend that assignment? You could say “I don't *have* to defend it—it just happens to be my credence.” But that seems about as unprincipled as looking at your sole source of infor-

For a defense of POI, see (White, 2009).

⁴ There are other kinds of cases that have been used to motivate imprecise credences. One motivation is the suggestion that credences don't obey **trichotomy**, which requires that for all propositions A and B , $c(A)$ is either greater than, less than, or equal to $c(B)$. (See e.g. (Schoenfield, 2012).) Another is the possibility of indeterminate chances, discussed in (Hájek & Smithson, 2012): if there are interval-valued chances, the Principal Principle seems to demand interval-valued credences. (I'll return to this argument in section 5.) Hájek & Smithson also suggest imprecise credences as a way of representing rational attitudes towards events with undefined expected value. Moss (2012) argues that imprecise credences provide a good way to model rational changes of heart (in the epistemic sense (if there is such a thing)). And there are still other motivations for the claim that ordinary agents are best modeled with imprecise credences, regardless of what rationality requires.

mation about the time, your digital clock, which tells that the time rounded off to the nearest minute is 4:03—and yet believing that the time is in fact 4:03 *and 36 seconds*. Granted, you may just happen to believe that; the point is that you have no business doing so. (Hájek & Smithson, 2012, 38–39)

[In Elga’s toothpaste/jellyfish case,] you may rest assured that your reluctance to have a settled opinion is appropriate. At best, having some exact real number assessment of the likelihood of more toothpaste would be a foolhardy response to your unspecific evidence. (Moss, 2012, 2)

The nonsharper’s position can be summarized with the following slogan: *unspecific or ambiguous evidence requires unspecific or ambiguous credences*.

1.2 What are imprecise credences?

Considerations like these suggest that sometimes (perhaps always) an agent’s credences should be *indeterminate* or *mushy* or *imprecise*, potentially over a wide interval. Some suggest that a rational agent’s doxastic attitudes are representable with an imprecise credence function, from propositions (sentences, events, ...) to lower and upper bounds, or to intervals within $[0, 1]$.⁵ A more sophisticated version of the view, defended by Joyce (2010), represents agents’ doxastic states with sets of precise probability functions. This is the version of the view that I’ll focus on.

I’ll use the following notation. C is the set of probability functions c that characterize an agent’s belief state; call C an agent’s “representor.” For representing imprecise credences toward propositions, we can say $C(A) = \{x : c(A) = x \text{ for some } c \in C\}$.

Various properties of an agent’s belief state are determined by whether the different probability functions in her representor have reached unanimity on that property. Some examples: an agent is more confident of A than of B iff for all $c \in C$, $c(A) > c(B)$. An agent is at least .7 confident in A iff $c(A) \geq .7$ for all $c \in C$. Her credence in A is .7 iff $c(A) = .7$ for all $c \in C$.

Similarly, if there’s unanimity among the credences in an agent’s representor, there are consequences for rational decision making: an agent is rationally required to choose an option ϕ over an option ψ if ϕ ’s expected utility is greater than ψ ’s relative to every $c \in C$. Beyond this sufficient condition, there’s some controversy

⁵ E.g. Kyburg (1983).

among nonsharppers about what practical rationality requires of agents with imprecise credences.⁶

2 Challenges for imprecise credences

The imprecise view has some appeal. But it also faces some major hurdles. In this section I'll discuss one general challenge for imprecise views, and two specific objections (one pragmatic and one epistemic). There's a natural strategy often mentioned by nonsharppers that might be able to handle both of these objections. In the next section (section 3), I offer the nonsharpper the most attractive version of this strategy for saving the imprecise view. But it's a trap: this version of the strategy would yield a version of the imprecise view that is indistinguishable from a particular kind of precise view. And this sort of precise view is well-equipped to handle ambiguous and unspecific evidence.

2.1 Fending off the permissivist alternative

There's one sort of precise view that can immediately accommodate the nonsharpper's starting intuition: that in cases like **toothpaste/jellyfish** and **coin of unknown bias**, the evidence doesn't uniquely support any precise credence function. *Permissivists* hold that there are bodies of evidence that don't single out one credence function as the unique rational credence function to adopt. But they don't conclude from this that we should have imprecise credences. Rather, the permissivist claims that there are multiple precise credence functions that are all equally rationally permissible.⁷

Permissivism is fairly controversial and faces some of its own challenges (see e.g. White 2005). So a satisfactory argument against nonsharppers can't simply end here. Furthermore, some of the objections to precise views that nonsharppers have put forward will apply equally to both precise permissivism and precise impermissivism. (I'll discuss these in section 5.)

Still, the possibility of permissivism is worth noting for two reasons. First, the path from ambiguous evidence to imprecise credences is not direct, and there are some well-known, respectable precise alternatives along the way. The objections to the imprecise view that I will give, and the alternative proposal I endorse, are equally available to both the permissivist and the impermissivist.

Second, the kinds of belief states that the imprecise view recommends are not clearly discernible from the kinds of belief states a permissivist recommends. For

⁶ See (Weatherson, 2008), (Joyce, 2010), (Williams, 2011), and (Moss, 2012).

⁷ See, for example, (Kelly, forthcoming), (Meacham, manuscript), and (Schoenfield, forthcoming).

example, when faced with a decision problem where values are held fixed, the imprecise view and the precise permissivist view can both allow that there are multiple permissible actions agents can take—multiple credences that would be permissible to bet in accordance with. So there is an important general challenge for the non-sharper: how can they explain what their view requires of our attitudes such that it's not simply a version of precise permissivism? The nonsharper might say: we demand credences that aren't fine-grained, and the permissivist allows fine-grained credences! But as a question about psychology, we need an account of what that amounts to. Functionalism or interpretivism will most naturally treat the two as interchangeable.

So: how can the imprecise view avoid collapsing into one form or another of the precise view? I'll leave this question open for the moment. But as I'll argue in section 3, in order for the imprecise view to give an adequate response to imprecision's greatest challenges, it may have to collapse into a precise view.

First, let me explain the two major challenges to the imprecise view: Elga's (2010) pragmatic challenge and White's (2009) epistemic challenge.

2.2 The pragmatic challenge

The pragmatic challenge for imprecise credences comes from Elga (2010). The argument is designed to show that there are certain kinds of rational constraints on decision making under uncertainty. The imprecise view faces difficulty in ensuring that rational agents with imprecise credences will satisfy these constraints.

Elga's challenge goes as follows: suppose you have an imprecise credence in some proposition A , say $C(A) = [.2, .8]$. We'll make the standard idealizing assumption that for you, value varies directly and linearly with dollars. You will be offered two bets about A , one very soon after the other (before you have time to receive new evidence or to change your priorities).

Bet 1 If A is true, you lose \$10. Otherwise you win \$15.

Bet 2 If A is true, you win \$15. Otherwise you lose \$10.

If you were pretty confident ($> .6$) that A was false, it would be rational for you to accept only bet 1; and if you were pretty confident that A was true ($> .6$), it would be rational for you to accept only bet 2. But since you're not confident of A or its negation, it seems like you should accept both bets; that way you'll receive a sure gain of \$5.

It is intuitively irrational to reject both bets, no matter what credences you have. The challenge for the nonsharper is to find some way to rule out the rationality of

rejecting both bets when your credences are imprecise. So far the nonsharper's only decision rule has been supervaluational: an agent is rationally required to choose an option over its alternatives if that option's expected utility is greater than its alternatives' relative to every $c \in C$. The expected value of each of our bets, considered in isolation, ranges from $-\$5$ to $\$10$, so neither is supervaluationally greater than the expected value of rejecting each bet ($\$0$). If this is a necessary and not just sufficient condition for being rationally required to choose an option—if the rule is *E-admissibility*—then in Elga's case, it'll be rationally permissible for you to reject both bets.⁸

Can the nonsharper give any good decision rule that prohibits rejecting both bets in this kind of case?⁹ If not, then imprecise credences make apparently irrational decisions permissible. After all, a pair of bets of this form can be constructed for imprecise credences in any proposition, with any size of interval. So it's not clear how the nonsharper can avoid permitting irrational decisions without demand-

⁸ Why? Well, consider t_1 , where you're offered bet 1, and t_2 , where you're offered bet 2. Whether you've accepted bet 1 or not, at t_2 it's permissible to accept or reject bet 2. (If you rejected bet 1, the expected value of accepting bet 2 ranges from $-\$5$ to $\$10$ and expected value of rejecting is $\$0$; so it's permissible to accept bet 2 and permissible to reject it. If you accepted bet 1, the expected value of accepting bet 2 is $\$5$ and the expected value of rejecting ranges from $-\$5$ to $\$10$, so it's permissible to accept bet 2 and permissible to reject it.)

From here, there's a quick possibility proof that in some circumstances E-admissibility permits rejecting both bets. Since at t_2 it's permissible to accept or reject bet 2, at t_1 you might be uncertain what you'll do in the future. Suppose you're .8 confident that you'll reject bet 2. There's a probability function in your representor according to which $c(A) = .8$. According to that probability function, the expected value of rejecting bet 1 is $\$2$ and the expected value of accepting bet 1 is $-\$3$. Since there's a probability function in your representor according to which rejecting bet 1 is better than accepting it, E-admissibility permits rejecting bet 1. So: in some circumstances, E-admissibility permits rejecting both bets.

⁹ Elga discusses some possible decision rules, but argues that most of them suffer substantial objections. For example, a more stringent decision rule might say: choosing an option is permissible only if that option maximizes utility according to every probability function in an agent's representor. But that rule entails that there will often be cases where there are *no* permissible options, including the case where you receive one of Elga's bets. For a case like that, it's uncontroversially absurd to think that there's nothing you can permissibly do. Other alternative decision rules, e.g. acting on the basis of the $c(A)$ at the midpoint of each $C(A)$, effectively collapse into precise views. Γ -Maximin—the rule according to which one should choose the option that has the greatest minimum expected value—prohibits rejecting both of Elga's bets (and requires accepting both). But that decision rule is unattractive for other reasons, including the fact that it sometimes requires turning down cost-free information. (See (Seidenfeld, 2004); thanks to Seidenfeld for discussion.) Still other rules, such as Weatherson's (2008) rule 'Caprice' and Williams's (2011) 'Randomize' rule, seem committed to the claim that what's rational for an agent to do depends not just on her credences and values, but also her past actions. This seems damningly akin to sunk cost reasoning. Unfortunately I don't have space to give these alternatives the attention they deserve.

ing that we always act on the basis of precise credences. And that, of course, teeters toward collapsing into the precise view.

2.3 The epistemic challenge

White (2009) delivers the epistemic challenge:

“Coin game: You haven’t a clue as to whether p . But you know that I know whether p . I agree to write ‘ p ’ on one side of a fair coin, and ‘ $\neg p$ ’ on the other, with whichever one is true going on the heads side. (I paint over the coin so that you can’t see which sides are heads and tails.) We toss the coin and observe that it happens to land on ‘ p ’.”
(175)

Let C be your credal state before the coin toss and let $C_{\cdot p'}$ be your credal state after. Our proposition p is selected such that, according to the nonsharper, $C(p) = [0, 1]$. And learning that a fair coin landed on the ‘ p ’ side has no effect on your credence in whether p is true, so $C_{\cdot p'}(p) = [0, 1]$. Ex hypothesi $p \equiv \text{heads}$, and so $C_{\cdot p'}(p) = C_{\cdot p'}(\text{heads})$.

So even though you know the coin is a fair coin, and seeing the land of the coin toss doesn’t tell you anything about whether the coin landed heads, the nonsharper says: the rational credence to have after seeing the coin toss is $C_{\cdot p'}(\text{heads}) = [0, 1]$.¹⁰ Your prior credence in heads was .5—it’s a fair coin—and seeing the result of the coin toss (‘ p ’) doesn’t tell you anything about whether heads. (After all, you have no clue whether ‘ p ’ is true.) But you should still update so that your new credence in heads is $[0, 1]$. That’s strange.

Furthermore, seeing the opposite result would also lead to the same spread of credences: $[0, 1]$. And so this case requires either a violation of Reflection or a violation of the Principal Principle. Reflection entails that if you know what credences your future (more informed, rational) self will have, you should have those credences now. So Reflection requires that since you know how the coin toss will affect your future credences, however the coin lands, you should have a *prior* credence $[0, 1]$ in heads—even though you know the chance of the coin landing heads is 50/50! So if you obey Reflection, you can’t obey the Principal Principle,

¹⁰ Why do $C_{\cdot p'}(p)$ and $C_{\cdot p'}(\text{heads})$ converge at $[0, 1]$ rather than at .5? Well, if $C_{\cdot p'}(p) = .5$ then for all $c \in C$, $c(p \mid 'p') = .5$. Symmetry surely requires the same update if the coin lands ‘ $\neg p$ ’, so for all $c \in C$, $c(p \mid '\neg p') = .5$. But then for all $c \in C$, $c(p \mid 'p') = c(p \mid '\neg p') = .5$. That entails the independence of ‘ p ’ and p , so for all $c \in C$, $c(p) = .5$. But this contradicts our initial assumption: $C(p) = [0, 1]$.

and vice versa.¹¹

This phenomenon is called *probabilistic dilation*. There are a few reasons why this is an unwelcome result. First, even if White’s coin game example is artificial, the features of the case that are sufficient for probabilistic dilation can appear in virtually any circumstances where you begin with a precise credence in some proposition. All that’s required is some possibility that that proposition is not probabilistically independent of some other proposition about which you have unspecific or ambiguous evidence.

Furthermore, it’s not a welcome result to end up with too many very spread out credences. One reason is that, at least with the sort of permissive decision theory that falls naturally out of the imprecise view, we’ll end up intuitively misdescribing a lot of decision problems: a wide range of intuitively irrational actions will be ruled rationally permissible.

Another reason is that a credence like $[0, 1]$ or $(0, 1)$ won’t allow for inductive learning, at least on the standard assumptions of the imprecise view (e.g. Joyce 2010). Take the **coin of unknown bias** example from above, where $C(\text{heads}) = C(\text{tails}) = (0, 1)$. Suppose you see the coin tossed two million times and it comes up heads one million times. Presumably for many probability functions in your representor, $c(\text{heads} \mid \frac{1M \text{ heads}}{2M \text{ tosses}})$ should be near .5. But, as Joyce acknowledges, there are at least two ways that your initial imprecision can force your credence to remain spread out at $(0, 1)$. First, your representor might include “pig-headed” probability functions that won’t move their probability assignment for heads out of some interval like $(.1, .2)$ no matter what you condition on. Second, “extremist” probability functions can prevent the total interval from moving. For each c that moves its credence in heads closer to .5 conditional on $\frac{1M \text{ heads}}{2M \text{ tosses}}$, there’s some more extreme function c^* such that $c(\text{heads}) = c^*(\text{heads} \mid \frac{1M \text{ heads}}{2M \text{ tosses}})$. So even if every credence function’s probability in heads moves toward .5, the posterior interval remains exactly where it was before the update.

So when your ignorance is extreme in this way, no inductive learning can take place.¹² And if White’s epistemic challenge is correct, then the nonsharper predicts that the circumstances that force you into this state of ignorance must be *very* widespread—even in cases where you have specific, unambiguous, and strong evidence about objective chances.

¹¹ Joyce argues there are ways of blocking the applicability of Reflection in this kind of case; see (Schoenfield, 2012) for a rebuttal to Joyce’s argument.

¹² See (Joyce, 2010), 290–291. Note that while probabilistic dilation is an inevitable outcome of any imprecise view, the induction-related challenge can be avoided by placing some (arguably ad hoc) constraints on which credence functions are allowable in an agent’s representor.

3 A poisoned pawn

A general strategy. How can the nonsharper respond to the pragmatic and epistemic challenges? A natural suggestion often comes up: both of these challenges seem solvable by **narrowing**. Narrowing amounts to sharpening $C(A)$ from some wide spread like $[0, 1]$ to something more narrow, maybe even some unique $c(A)$. Elga (2010) and Joyce (2010) both discuss this option. The narrowing strategy can be invoked for both the epistemic and pragmatic challenges.

For the pragmatic challenge: narrowing one's credences to a singleton subset of C will certainly guarantee that a rational agent presented with any pair of bets like Elga's will not reject both.¹³ After all, either $c(A)$ is greater than .4 or it's not. If $c(A) > .4$, she will accept bet 2. And if $c(A) \leq .4$ —indeed, if $c(A)$ is anywhere below .6—then she will accept bet 1. So whatever the rational agent's credence in A is, she will accept at least one of the bets. And even narrowing to a non-singleton (i.e. imprecise) subset of C , faced with a particular bet, can avoid the challenge. For example, $C(A) = (.4, .6)$ ensures that you should accept both bets: since all probability functions in C assign A credence greater than .4, accepting bet 2 has greater expected value than rejecting it according to all $c \in C$. And since they all assign A credence less than .6, accepting bet 1 has higher expected value than rejecting it according to all $c \in C$.

For the epistemic challenge: one way to narrow one's credence so that inductive learning is possible is by whittling off extremist and pig-headed credence functions. And of course, narrowing to a unique non-pig-headed credence will allow for inductive learning.

Some obstacles. If narrowing is understood as effecting a change in one's epistemic state, in order to achieve some goal (making a decision, performing induction), then there's an immediate objection. Narrowing can't be rationally permissible, because it involves changes in one's epistemic state without changes in one's evidence. For the pragmatic case, though, Joyce (2010) argues that this objection can be avoided. Narrowing can be thought of not as a change in an agent's epistemic state, but a change in how the agent's epistemic state is linked to how she chooses actions to carry out. Pragmatic narrowing has no effect on an agent's epistemic state; the narrowed down set of credence functions doesn't represent the agent's beliefs. It just represents the subset of her representor that is relevant for making the decision at hand.

Of course, this response is not available for the epistemic case. At best this would allow for a pretense of learning. But for the epistemic case, we might not

¹³ Assuming her decision doesn't change the probability of the hypothesis in question.

think of narrowing as a diachronic adjustment of one's credal state. It might just be that some sort of narrowing of C has been rationally required all along. So some sort of narrowing strategy might well be valuable for both cases separately. And ideally, we could put the strategy to work for both in some unified way.

A deeper worry: how can narrowing be accomplished in a non-ad-hoc way? What kinds of plausible epistemic or pragmatic considerations would favor ruling some credence functions in and ruling others out? For example, in the epistemic case, we might want to shave off extremist and pig-headed credence functions. But to avail themselves of that strategy, nonsharppers need to offer some epistemic justification for doing so.

A sophisticated narrowing strategy. What would justify eliminating pig-headed and extremist credences? Let me offer a hypothesis: a rational agent might be comparatively confident that these sorts of narrowings are less likely to be reasonable to adopt, even for pragmatic purposes, than credence functions that aren't extremist or pig-headed.

Of course, to say a sharp credence function is *reasonable to adopt* for narrowing won't, on the imprecise view, be the same as saying the credence function is *rational to have*. According to nonsharppers, no sharp credence function is rational, at least relative to the kinds of evidence we typically face. And it's a bit tricky to say what this second normative notion, "reasonability," is.

But arguably this notion already figures into nonsharppers' epistemology. After all, most nonsharppers don't hold that our credence should be $[0, 1]$ for every proposition that we have any attitude toward. Consider some proposition that, by nonsharper lights, doesn't require a maximally spread out credence relative to an agent's evidence. Why doesn't her representor include probability functions that assign extreme values to that proposition?

It must be that the prior probability functions in an agent's representor have to meet some constraints. Some probability functions are rationally inappropriate to include in our credal states. Joyce, in introducing the notion of a representor, writes: "Elements of C are, intuitively, probability functions that the believer takes to be compatible with her total evidence" (288). Joyce certainly can't mean that these probability functions are rationally permissible for the agent to have as her credence function. On Joyce's view, none of the elements of C is rationally permissible to have as one's sole credence function; they're all precise. So he must mean that these probability functions all have some other sort of epistemic credentials. We can leave precisely what this amounts to a black box, for a moment, and explore what possibilities it opens up.

How does this help? Suppose it's correct that for some epistemic reasons, there

are probability functions that cannot be included in a rational agent's representor, C . It seems plausible that of the remaining probability functions in C , some might have better or worse epistemic credentials—that is, are more reasonable—than others. For example, in the **coin of unknown bias** case, even if $C(A) = (0, 1)$, it's hard to believe a rational agent would bet at odds other than .5. The probability functions according to which $c(\text{heads}) = .5$ seem plausibly more reasonable to act on than probability functions according to which $c(\text{heads}) = .999$ or $= .0001$.

So we can wrap our heads around the idea that some probability functions in C deserve more weight in guiding our decisions than others. But that doesn't necessarily justify eliminating those probability functions with worse credentials from C . After all, there will be borderline cases. Consider, for example, which "extremist" credences could be eliminated.

The possibility of borderline cases generates two worries.¹⁴ First, narrowing by simply drawing a boundary between very similar cases seems arbitrary: it's unclear what would justify drawing the narrowing's boundary between some particular pair of probability functions, rather than between some other pair. Second, we end up predicting a big difference in the relevance of two probability functions in determining rational actions, even though there's not a big difference in how reasonable they seem to be.

There's a natural alternative: instead of simply ruling some credence functions in and ruling others out, we can impose a weighting over $c \in C$ such that all probability functions in C carry some weight, but some probability functions carry more weight than others in determining what choices are rational. What weighting? Plausibly: the agent's rational credence that c is a reasonable narrowing to adopt. For example, pig-headed credence functions seem clearly less likely to be reasonable narrowings to act on than credence functions that are responsive to inductive evidence.

A weighting over probability functions in an agent's representor can be used to determine a narrowed down credence for whatever proposition is in question. The narrowing to adopt in A could be some sort of weighted average of each $c(A)$ such that $c \in C$.

This proposal yields a narrowing strategy that has natural epistemic motivations. Acting on a narrowing based on a weighting doesn't require ad hoc choices in the way that simply eliminating probability functions does. And furthermore, this guarantees that the agent's multiple credence functions aren't idle in determin-

¹⁴ Alan Hájek (2012) distinguishes these two types of concerns, both of which often feature in worries throughout philosophy, about the arbitrariness involved in drawing sharp distinctions between borderline cases.

ing which narrowing she adopts. None simply gets thrown out.

So, have we solved the problems for the nonsharper?

The trap. The kind of imprecise view that results from this proposal is really just a complicated notational variant of a precise view. Instead of having multiple probability functions, the rational agent has a single probability function that gives positive probability to the rationality of other credence functions. In other words, this precise credence function includes uncertainty about its own rationality and about the rationality of its competitors. The weighting of $c \in C$ are just the agent's credences in each c 's rationality.

Now, there might be other strategies that the nonsharper could adopt for coping with the pragmatic and epistemic challenges. But I draw attention to this strategy for two reasons.

First, I think it's a reasonably attractive strategy the nonsharper can adopt for addressing these challenges. It's certainly the only narrowing strategy I know of that provides a motivation for how and why the narrowing can take place.

Second, it brings into focus an observation that I think is important in this debate, and that has been so far widely ignored: *higher-order credences and higher-order uncertainty can play the role that imprecise credences were designed for*. Indeed, I'm going to try to convince you that they can play that role even better than imprecise credences can.

4 The precise alternative

Before showing how this sort of precise view can undermine the motivations for going imprecise, let's see what exactly the view amounts to.

We can expect rational agents to have higher-order credences of various kinds: in particular, credences about whether or not their own credence functions are rational, and credences about whether other possible credence functions are rational. I use the phrase "higher-order credences" broadly to mean credences about credences. Some philosophers reserve the phrase for credences about what credences one has (for characterizing some sort of introspective uncertainty). That's not what I'll be talking about.

It's compatible with even ideal rationality for an agent to be uncertain about whether her credences are rational.¹⁵ An ideally rational agent can even be some-

¹⁵ Uncertainty about whether one's credences are rational is a form of *normative* uncertainty. Since normative truths are generally held to be metaphysically necessary—true in all possible worlds—we need to take care in modeling normative uncertainty, in order to ensure that normative truths aren't automatically assigned probability 1. My own preference is to use an enriched possible worlds

what confident that her own credences are not rational.¹⁶ For example: if a rational agent is faced with good evidence that she tends to be overconfident about a certain topic and good evidence that she tends to be underconfident about that same topic, she may doubt that the credence she has is *exactly* the credence that's warranted by her evidence. So a rational agent may not know whether she's responded correctly to her evidence. There are also some cases, too complex to discuss here, where an ideally rational agent might simply not be in a position to know what her evidence *is*, and therefore be uncertain whether her credence is warranted by the evidence.¹⁷

Nonsharper hold that there are some bodies of evidence that are unspecific or ambiguous. These bodies of evidence, according to the nonsharper, rationally require that agents adopt a state that encompasses *all* credence functions compatible (in some sense) with the evidence. On the precise view that I'm advocating, if there really is ambiguous or unspecific evidence, then if faced with these bodies of evidence, rational agents will simply be *uncertain* what credences it is rational to have. That's compatible with continuing to have precise credences. Instead of attributing all of the candidate probability functions to the agent, we push this set of probability functions up a level, into the contents of the agent's higher order credences.

A caveat: it's compatible with the kind of view I'm defending that there are no such bodies of evidence. It might be that every body of evidence not only supports precise credences, but supports certainty in the rationality of just those precise credences.

Here is my claim: if there really are bodies of ambiguous or unspecific evidence, then these bodies of evidence support higher-order uncertainty.¹⁸ Elga's **toothpaste/jellyfish** case is a promising candidate: when you're met with such an odd body of evidence, you should be uncertain what credence would be rational to have. And indeed, I think the nonsharper should agree with this point. What would justify a spread-out credence like $c(\text{TOOTHPASTE}) = [.2, .8]$ over $[.1999, .7999]$?

I also claim that once we take into account higher-order uncertainty, we'll see that first-order imprecision is unmotivated. For example, consider the argument Joyce (2010) uses to support imprecise credences in the **coin of unknown bias**

probability space, as in (Gibbard, 1990). Instead of defining probability as a measure over the space of possible worlds, we define it over the space of world–norm pairs, where “norms” are maximally specific normative theories.

¹⁶ See e.g. (Elga, 2008), (Christensen, 2010), and (Elga, 2012).

¹⁷ See (Williamson, 2007), (Christensen, 2010), and (Elga, 2012).

¹⁸ Note that they provide only a sufficient condition; there could be unambiguous, specific bodies of evidence that also support higher-order uncertainty.

case:

f_U [the probability density function determined by POI] commits you to thinking that in a hundred independent tosses of the [coin of unknown bias] the chances of [heads] coming up fewer than 17 times is exactly 17/101, just a smidgen ($= 1/606$) more probable than rolling an ace with a fair die. Do you really think that your evidence justifies such a specific probability assignment? Do you really think, e.g., that you know enough about your situation to conclude that it would be an unequivocal mistake to let \$100 ride on a fair die coming up one rather than on seeing fewer than seventeen [heads] in a hundred tosses? (284)

Answer: No. If the **coin of unknown bias** case is indeed a case of unspecific or ambiguous evidence, then I'm *uncertain* about whether my evidence justifies this probability assignment; I'm uncertain about what credences are rational. And so I'm uncertain about whether it would be an unequivocal mistake to bet in this way. After all, whether it's an unequivocal mistake is determined by what credences are *rational*, not whatever credences I happen to have.

But if I'm uncertain about which credences are rational, there's no reason why I should adopt *all* of the candidates. (If I'm uncertain whether to believe these socks are black or to believe they're navy, should I adopt both beliefs?)

Of course, this only tells us something about the higher-order credences that are rational to have in light of unspecific or ambiguous evidence: that they can reflect uncertainty about what's rational, and that they're compatible with sharp first-order credences. One might ask: *which* sharp first-order credences are rational? After all, the nonsharper's initial challenge to the sharper was to name *any* first order credences in the **toothpaste/jellyfish** case that could seem rationally permissible.

But the kind of view I'm advocating *shouldn't* offer an answer to that question. After all, didn't I just say that we should be uncertain? It would be a pragmatic contradiction to go on and specify what sorts of first-order credences are rational.¹⁹

¹⁹ Unofficially, I can mention some possible constraints on rational lower-order credences. What's been said so far has been neutral about whether there are *level-bridging norms*: norms that constrain what combinations of lower and higher-order credences are rational. But a level-bridging response, involving something like the principle of Rational Reflection, is a live possibility. (See (Christensen, 2010) and (Elga, 2012) for a important refinement of the principle.) According to this principle, our rational first-order credences should be a weighted average of the credences we think might be rational (conditional on their own rationality), weighted by our credence in each that it is rational. Formally, Elga's version of Rational Reflection says: where p_i is a candidate rational credence function, $c(A \mid p_i \text{ is ideal}) = p_i(A \mid p_i \text{ is ideal})$. This principle determines what precise probabilities an agent

So now we have a precise view that shows the same sensitivity to ambiguous and unspecific evidence as the imprecise view. In effect, it does all the work that the imprecise view was designed for, without facing the same challenges. So are there any reasons left to go imprecise? In the remainder of this paper, I'm going to argue that there aren't.

5 The showdown

Here is the dialectic so far. There is a lot of pressure on nonsharper to move in the direction of precision. The pragmatic and epistemic challenges both push in that direction. But if rational agents must act as though they have precise credences, then on the widely presupposed *interpretivist* view of credences—that whatever credences the agent has are those that best explain and rationalize her behavioral dispositions—then the game is up. As long as imprecise credences don't play a role in explaining and rationalizing the agent's behavior, they're a useless complication.²⁰

But the nonsharper might bite the bullet and reject interpretivism. Even if rational agents are disposed to *act* as though they have precise credences (in all possible situations!), the nonsharper might claim, epistemic rationality nevertheless demands that they *have* imprecise credences. These imprecise credences might play no role in determining behavior. Still, the nonsharper might say, practical and epistemic norms impose different but compatible requirements. Practical norms might require acting on a precise credence function, but epistemic norms require having imprecise credences.²¹

should have when she has rational higher-order uncertainty. The Christensen/Elga principle might not be the last word, but it's an attractive hypothesis. Note, however, that a principle like this won't provide a recipe to check whether your credences are rational: whatever second-order credences you have, you'll also be uncertain about whether your second-order credences are the rational ones to have, and so on. And so, again, it would be inconsistent with the view I'm offering to provide a response to the question of which sharp first-order credences are rational.

²⁰ Hájek & Smithson (2012) argue that interpretivism directly favors modeling even ideal agents with imprecise credences. After all, a finite agent's dispositions won't determine a unique probability function/utility function pair that can characterize her behavioral dispositions. And this just means that all of the probability/utility pairs that characterize the agent are equally accurate. So, doesn't interpretivism entail at least the permissibility of imprecise credences? I find this argument compelling. But it doesn't tell us anything about epistemic norms (beyond some application of *ought* implies *can*, which is always on questionable footing in epistemology). It doesn't suggest that evidence ever makes it rationally required to have imprecise credences. And so this argument doesn't take sides between the imprecise and precise views that I'm concerned with.

²¹ Note that this also requires biting the bullet on the epistemic challenge.

This bullet might be worth biting if we had good evidence that epistemic norms in fact do require having imprecise credences. Then the nonsharper would be able to escape the charge, from section 3, that any adequate narrowing strategy collapses their view into the precise view (though again, at the cost of rejecting interpretivism). So the big question is: *Is there any good motivation for the claim that epistemic norms require imprecise credences?*

I'm going to argue that the answer is *no*. Any good motivation for going imprecise is at least equally good, and typically better, motivation for going precise and higher-order. In this section, I'll consider a series of progressive refinements of the hypothesis that imprecise evidence mandates imprecise credences—each a bit more sophisticated than the last. I'll explain how each motivation can be accommodated by a precise view that allows for higher-order uncertainty. The list can't be exhaustive, of course. But it will show that (to borrow from the late linguist Tanya Reinhart) the imprecise view has a dangerously unfavorable ratio of solutions to problems.

5.1 Nonsharper claim #1: Precise credences should reflect known chances

In motivating their position, nonsharper often presuppose that without knowledge of objective chances, it's inappropriate to have precise credences. Here's an example:

A...proponent of precise credences...will say that you should have some sharp values or other for [your credence in drawing a particular kind of ball from an urn], thereby committing yourself...to a definite view about the **relative proportions** of balls in your urn... Postulating sharp values for [your credences] under such conditions amounts to pulling **statistical correlations** out of thin air. (Joyce, 2010, 287, emphasis added)

Or again:

f_U commits you to thinking that in a hundred independent tosses of the coin [of unknown bias] the **chances** of [heads] coming up fewer than 17 times is exactly 17/101, just a smidgen (= 1/606) more probable than rolling an ace with a fair die. (Joyce, 2010, 284, emphasis added)

It's difficult to see what exactly Joyce is suggesting. On a naïve interpretation, he seems to be endorsing the following principle:

CREDENCE/CHANCE: having credence n in A is the same state as, or otherwise necessitates, having credence ≈ 1 that the objective chance of A is (or at some prior time was) n .²²

A somewhat flat-footed objection seems sufficient here: one state is a partial belief, the content of which isn't about chance. The other is a full belief about chance. So surely they are not the same state.

More generally: whether someone takes a definite stance isn't the kind of thing that can be read locally off of her credence in A . There are *global* features of an agent's belief state that determine whether that credence reflects some kind of definite stance, like a belief about chance, or whether it simply reflects a state of uncertainty.

For example, in the **coin of unknown bias** case, someone whose first-order credence in HEADS is .5 on the basis of applying the principle of indifference will have different attitudes from someone who believes that the objective chance of HEADS is .5. The two will naturally have different introspective beliefs and different beliefs about chance. The former can confidently claim: "I don't have any idea what the objective chance of HEADS is"; "I doubt the chance is .5"; etc. Neither is rationally compatible with taking a definite position that the chance of HEADS is .5.

The agent who is uncertain about chance will exhibit other global differences in her credal state from the agent with a firm stance on chance. A credence, $c(A) = n$, doesn't always encode the same degree of *resiliency* relative to possible new evidence. The resiliency of a credence is the degree to which it is stable in light of new evidence.²³ When an agent's credence in A is n because she believes the chance of A is n , that credence is much more stubbornly fixed at or close to n . Credences grounded in the principle of indifference, in ignorance of objective chances, are much less resilient in the face of new evidence.²⁴ For example, if your .5 credence is grounded in the principle of indifference and then you learn that the last three tosses of the coin have all landed heads, you'll substantially revise your credence that the next coin will land heads. (After all, three heads is some evidence that the coin is biased toward heads.) But if your .5 credence comes from the knowledge that the chance is .5, then your credence shouldn't change in response to this

²² This is a slightly stronger version of what White (2009) calls the *Chance Grounding Thesis*, which he attributes to a certain kind of nonsharper: "Only on the basis of known chances can one legitimately have sharp credences. Otherwise one's spread of credence should cover the range of chance hypotheses left open by your evidence" (174).

²³ See (Skyrms, 1977).

²⁴ See also (White, 2009, 162–164).

evidence.

In short: the complaint against precise credences that Joyce seems to be offering in the passages quoted above hinges on a false assumption: that having a precise credence in a hypothesis *A* requires taking a definite view about chance. Whether or not an agent takes a definite view about the chance of *A* isn't determined locally by the precision of her credence in *A*. It depends on other properties of her credal state, which can vary independently of her credence. And so having a precise credence is compatible with having no definite views about chance.

5.2 Nonsharper claim #2: Precise credences are “too informative”

The second motivation for imprecise credences is a generalization of the first. Even if precise credences don't encode full beliefs about chances, they still encode something that shouldn't be encoded: they're still too unambiguous and specific a response to ambiguous or unspecific evidence.

Even if one grants that the uniform density is the *least* informative sharp credence function consistent with your evidence, it is still *very* informative. Adopting it amounts to pretending that you have lots and lots of information that you simply don't have. (Joyce, 2010, 284)

How would you defend that assignment? You could say “I don't *have* to defend it—it just happens to be my credence.” But that seems about as unprincipled as looking at your sole source of information about the time, your digital clock, which tells that the time rounded off to the nearest minute is 4:03—and yet believing that the time is in fact 4:03 *and 36 seconds*. Granted, you may just happen to believe that; the point is that you have no business doing so. (Hájek & Smithson, 2012, 38–39)

Something of this sort seems to underpin a lot of the arguments for imprecise credences. But is this right?

Well, there's a clear sense in which specifying a set of probability functions can be less informative than specifying a unique probability function. In science and statistics, imprecise probabilities are used in cases where, because there is little information, only a partial specification of probability can be given. So, when the chances of a set of events aren't fully known, imprecise probabilities are useful for representing both what chance information is available and the ways in which it is limited. Imprecise probabilities are less informative about objective chances.

But this only lends support to using a certain kind of mathematical apparatus to represent chance. It certainly doesn't suggest that our mental states should be imprecise. After all, scientists and statisticians almost all assume that there are precise objective chances; they're just uncertain which probability function correctly represents them, and so can only give an imprecise specification of chance. And so analogously, suppose we can only give an imprecise specification of the rational epistemic probabilities, or of the degrees to which the evidence confirms a hypothesis. Then, by analogy, we should be *uncertain* which probability function is rational to adopt, or about the degree of evidential confirmation. But that's not the nonsharper view. That's *my* view.

So, the analogy with imprecise probabilities in science, statistics, and probability theory does not support the nonsharper view. It's certainly true that imprecise credences encode less information than precise credences, for the same reason that, in a possible worlds framework, a non-singleton set of worlds encodes less information than a single world. But the real questions are:

- (1) what kind of information is encoded, and
- (2) is it problematic or irrational to encode that information?

The answers: (1) information about agents' mental states, and (2) no.

An ascription of a precise credence function is more informative than an ascription of a set of credence functions. After all, if you tell me that an agent has a credence $[.2, .7]$ in A , I know less about what bets she'll be inclined to accept than if you tell me that she has credence .34.

But it's not more informative about things like coin tosses or their objective chance. Instead, it's more informative about the psychology and dispositions of an agent. This is third-personal information offered by the theorist about an agent's attitudes, not information in the contents of agent's first-order attitudes.²⁵ Precise credences are unambiguous and specific about agents' doxastic states. They tell us, for example, precisely how a rational agent will bet once we've fixed her utilities. But why would there be anything wrong with being informative or unambiguous or specific in this way?

It's uncontroversial that in a case like **coin of unknown bias**, an agent should not presume to have information about how the coin will land, given how little

²⁵ Of course, the rational agent may ascribe herself precise or imprecise credences and so occupy the theorist's position. But in doing so, the comparative informativeness in her ascription of precise credences is informativeness about her own psychological states, not about how coin-tosses might turn out.

evidence she has. In that sense, since she has almost no information, the less information *about coin tosses* that she takes herself to have, the better. But that doesn't mean that the less information the theorist has about the rational agent's *mental states*, the better. And that is what is represented in assigning a precise credence function.

After all: the rational agent is stipulated to have limited information available to her, and so her beliefs should reflect that fact. She should be uncertain about the coin's chances of landing heads (just like the scientists and statisticians). But there is no similar stipulation that the theorist has limited information in characterizing the rational agent. So there's just no reason why the theorist's assignments of credences to the agent should be uninformative.

Objection 1. In the coin of unknown bias case, if $c(\text{HEADS}) = .5$, then you are taking a specific attitude toward how likely it is that the coin lands heads. You think the coin is .5 likely to land heads. That *is* information about the coin that the agent is presuming to have.

Reply. What does it mean, in this context, to say *The coin is .5 likely to land heads*? It doesn't mean that you think the *chance* of the coin landing heads is .5; you don't know whether that's true. It doesn't even mean that you think the *evidential probability* of the coin landing heads is .5; you can be uncertain about that as well. "It's .5 likely that heads" arguably doesn't express a belief at all. It just expresses the state of having .5 credence in the coin's landing heads.²⁶ But then the .5 part doesn't tell us anything about the coin. It just expresses some aspect of your psychological state.

Objection 2. If the evidence for a proposition *A* is genuinely imprecise, then there is some sense in which adopting a precise credence in *A* means not withholding judgment where you really ought to.

Reply. If my credence in *A* is not close to 0 or 1, then I'm withholding judgment about whether *A*. That's just what withholding judgment is. The nonsharper seems to think that for some reason I should double down and withhold judgment *again*. Why? It can't be because I'm not withholding judgment about what the evidence supports; higher-order uncertainty takes care of that. If my credence in the proposition *the evidence supports my credence in A* is also not close to 0 or 1, then I'm clearly withholding judgment about what the evidence supports.

In short: there's just no reason to believe the slogan that ambiguous or unspecific evidence requires ambiguous or unspecific credences. Why should the *attitude* be confusing or messy just because the evidence is? (If the evidence is unimpres-

²⁶ Cf. Yalcin (forthcoming).

sive, that doesn't mean our credences should be unimpressive.) What is true is that ambiguous or unspecific evidence should be reflected in one's beliefs, somehow or other. But that might amount to simply believing that the evidence is ambiguous and unspecific, being uncertain what to believe, having non-resilient credences, and so on. And all of these are naturally represented within the precise model.

Finally, let's consider one more refinement of this objection, one that can give some argument for the hypothesis that imprecise evidence requires imprecise credences.

5.3 Nonsharper claim #3: Imprecise confirmation requires imprecise credences

A different form of argument for imprecise credences involves the following two premises:

IMPRECISE CONFIRMATION The confirmation relation between bodies of evidence and propositions is imprecise.

STRICT EVIDENTIALISM Your credences should represent only what your evidence confirms.

These two claims might be thought to entail the imprecise view.²⁷

According to nonsharppers, the first claim has strong intuitive appeal. It says that, for some bodies of evidence and some propositions, there is no unique precise degree to which the evidence supports the proposition. Rather, there are multiple equally good precise degrees of support that could be used to relate bodies of evidence to propositions. This, in itself, is not a claim about rational credence, any more than claims about entailment relations are claims about rational belief. So in spite of appearances, this is not simply a denial of the precise view, though the two are tightly related.

In conjunction with STRICT EVIDENTIALISM, though, it might seem straightforwardly impossible for the sharper to accommodate IMPRECISE CONFIRMATION. Of course, some sharpeners consider it no cost at all to reject IMPRECISE CONFIRMATION. They might have considered this a fundamental element of the sharper view, not some extra bullet that sharpeners have to bite.

But whether rejecting IMPRECISE CONFIRMATION is a bullet or not, sharpeners don't have to bite it. The conjunction of IMPRECISE CONFIRMATION and STRICT EVIDENTIALISM is compatible with the precise view.

²⁷ Thanks to Wolfgang Schwarz, Rachael Briggs, and Alan Hájek for pressing me on this objection.

It's clear that IMPRECISE CONFIRMATION is compatible with one form of the precise view, namely permissivism. If there are a number of probability functions that each capture equally well what the evidence confirms, then precise permissivists can simply say: any of them is permissible to adopt as a credence function. Permissivism is practically designed to accommodate IMPRECISE CONFIRMATION.

Of course, some nonsharppers might think that adopting a precise credence function on its own would amount to violating STRICT EVIDENTIALISM. But this suggestion was based on the assumption that precise credences are somehow inappropriately informative, or involve failing to withhold judgment when judgment should be withheld. In the last two subsections of this paper, I've argued that this assumption is false. Precise permissivism is compatible with both claims.

Perhaps more surprisingly, precise impermissivism is also compatible with both claims. If IMPRECISE CONFIRMATION is true, then some bodies of evidence fail to determine a unique credence that's rational in each proposition. And so epistemic norms sometimes don't place a determinate constraint on which probability function is rational to adopt. But this doesn't entail that the epistemic norms require adopting multiple probability functions, as the nonsharper suggests. It might just be that in light of some bodies of evidence, *epistemic norms place only an indeterminate constraint on our credences*.

Suppose this is right: when our evidence is ambiguous or unspecific, it's indeterminate what rationality requires of us. This is compatible with the precise view: it could be supervalueationally true that our credences must be precise. Moreover, this is compatible with impermissivism: it could be supervalueationally true that it's not the case that more than one credence function is permissible.

How could it be indeterminate what rationality requires of us? There are cases where morality and other sorts of norms don't place fully determinate constraints on us. Here is a (somewhat idealized) example. When I'm grading, I may be obligated to give As to excellent papers, A-s to great but not truly excellent papers, B+s to good but not great papers, and so on. Suppose some paper I receive is a borderline case of a great paper: it's not determinately great and not determinately not great. And so here, it seems like I'm not determinately obligated to assign a B+, nor am I determinately obligated to assign an A-. There's an indeterminacy in my obligations. But this clearly doesn't mean that I have some obligation to mark the student's paper with some sort of squiggle such that it's indeterminate whether the squiggle is a B+ or an A-.²⁸

²⁸ Roger White suggested a similar example in personal communication.

The upshot is clear: *Indeterminacy in obligations doesn't entail an obligation to indeterminacy*.²⁹ In this case, obviously I'm obligated to give a precise grade, even if it's indeterminate which precise grade is required.

It might be protested that if the norms don't fully determine my obligations, then it must be that either grade is permissible. But according to the (idealized) setup, I'm obligated to give an A- iff a paper is great but not excellent and to give a B+ iff a paper is good but not great. This paper is either great or good but not great. So either I'm obligated to give an A- or I'm obligated to give a B+. The indeterminacy doesn't imply that the norms are overturned and neither disjunct is true. If anything, it implies that it's indeterminate whether a B+ is permissible or an A- is permissible. Analogously: if IMPRECISE CONFIRMATION is correct, then it might not be true of any credence function that it's determinately required (in light of some evidence). But that doesn't mean that more than one probability function is determinately permissible.

Furthermore, if both grades were permissible, then the choice between them would be arbitrary (relative to the norms of grading). And we could imagine it to be a further norm of grading that one never assign grades arbitrarily. So the norms could be overtly impermissive. Then there's no getting around the fact that, according to the norms of grading we've stipulated, I have no choice but to take some action that isn't determinately permissible.

²⁹ This point extends to another argument that has been given for imprecise credences. According to Hájek & Smithson (2012), there could be *indeterminate chances*, so that some event E 's chance might be indeterminate—not merely unknown—over some interval like $[.2, .5]$. This might be the case if the relative frequency of some event-type is at some times .27, at others .49, etc.—changing in unpredictable ways, forever, such that there is no precise limiting relative frequency. Hájek & Smithson argue that the possibility of indeterminate objective chances, combined with the following natural generalization of Lewis's Principal Principle, yields the result that it is rationally required to have imprecise or (to use their preferred term) *indeterminate* credences.

PP* Rational credences are such that $C(A \mid Ch(A) = [n, m]) = [n, m]$ (if there's no inadmissible evidence).

But there are other possible generalizations of the Principal Principle that are equally natural, e.g. PP†:

PP† Rational credences are such that $C(A \mid Ch(A) = [n, m]) \in [n, m]$ (if there's no inadmissible evidence):

The original Principal Principle is basically a special case of both. (Note that PP† only states a necessary condition on rational credences and not a sufficient one. So it isn't necessarily a permissive principle.) Hájek & Smithson don't address this alternative, but it seems to me perfectly adequate for the sharper to use for constraining credences in the face of indeterminate chances. Again, we cannot assume that indeterminacy in chances requires us to have indeterminate credences.

Analogously: even if epistemic norms underdetermine what credences are rational, it might still be the case that we're epistemically and rationally required to adopt a precise credence function, and furthermore that impermissivism is true. This might seem puzzling: if the evidence genuinely underdetermines which credences to have, then how could precise impermissivism be true? Well, it might be an epistemic norm that we reject permissivism. (There's some motivation for this: there's something intuitively problematic about having a credence function which you take to be rational, but thinking that you could just as rationally have had a different credence function.) If this is so, no fully rational credence function assigns nonnegligible credence to the possibility that multiple credence functions are appropriate responses to a single body of evidence.³⁰ So precise impermissivism, like precise permissivism, has no fundamental problem with accommodating IMPRECISE CONFIRMATION.

One concern I've often heard is that there's some analogy between the view I defend and epistemicism about vagueness. Epistemicism is, of course, the view that vague predicates have perfectly sharp extensions. We just don't what those extensions are; and this ignorance explains away the appearance of indeterminacy. One might think that the impermissive version of my view amounts to something like an epistemicism about ambiguous evidence. Instead of allowing for the possibility of genuine indeterminacy, the thought goes, my view suggests we might simply not know what sharp credences are warranted. Still, though, the credence that's required is perfectly sharp.

But a precise view that countenances genuine indeterminacy—that is, indeterminacy that isn't merely epistemic—is fundamentally different from epistemicism about vagueness. And allowing for indeterminate epistemic requirements, and so IMPRECISE CONFIRMATION, is clearly allowing for genuine indeterminacy. The supervenient story I offered above quite closely analogous to one of epistemicism's major opponents, supervenientism. The supervenientist about vagueness holds that there is determinately a sharp cut-off point between non-bald and bald; it just isn't determinate where that cut-off point is. Similarly, the precise impermissivist who accepts IMPRECISE CONFIRMATION accepts that for any body of evidence, there is determinately a precise credence function one ought to have in light of that evidence; it's just indeterminate what that precise credence function is.

³⁰ Of course, that's compatible with permissivism's being true; maybe we're epistemically required to accept a falsehood. But if impermissivists are right that it's a norm of rationality to reject permissivism, then they must accept that the norms of rationality apply to them, and so reject permissivism. The degree to which you are a realist or antirealist about epistemic norms will probably affect how problematic you find this possibility.

6 Conclusion

The nonsharper claims that imprecise evidence requires imprecise credences. I've argued that this is false: imprecise (ambiguous, nonspecific) evidence can place special constraints on our attitudes, but not by requiring our attitudes to be imprecise. The nonsharper's view rests on the assumption that having imprecise credences is the only way to exhibit certain sorts of uncertainty: uncertainty about chance (objective probability), about rational requirements (evidential probability), or about confirmation (logical probability). I've argued that these sorts of uncertainty can naturally be captured within the precise framework. All we need are higher-order probabilities: subjective probability about other forms of probability, like chance and ideally rational probability.

The kind of precise view I defend can accommodate all the intuitions that were taken to motivate the imprecise view. So what else does going imprecise gain us? As far as I can tell, only vulnerability to plainly irrational diachronic decision behavior and an inability to reliably use Reflection or to reason by induction.³¹ Better to drop the imprecision and stick with old-fashioned precise probabilities.

³¹ In other words, the pragmatic and epistemic challenges from section 2.

Chapter 2. Don't Stop Believing

Epistemic rationality requires two kinds of coherence. Broadly speaking, an agent's beliefs must fit well together at a time, and also fit well together over time. At any particular time, we should avoid believing contradictions, believe the consequences of our beliefs, and so on. And over time, we should respect the evidence we've received and adapt our beliefs to new evidence.

The traditional Bayesian picture of epistemic rationality is simply the conjunction of a synchronic claim and a diachronic claim:

Synchronic coherence: Rational belief states form a probability function and are rationalized by one's evidence.

Diachronic coherence: Rational belief states evolve by retaining old certainties and conditioning on new evidence.

Recently, however, a number of philosophers have pushed for the abandonment of diachronic norms. Norms like Conditionalization, that have historically been understood as constraints on beliefs at different times, have been reinterpreted as purely synchronic constraints. According to this view, the norms of rationality, practical or epistemic, apply only to time-slices of individuals.

I want to resist this movement. I'll argue for the following claim:

Diachronic Rationality: There are diachronic norms of epistemic rationality.

The problem that the opponent of diachronic rationality poses is this: diachronic norms of epistemic rationality are in tension with *epistemic internalism*. Epistemic internalism, in its most generic form, is the view that whether or not you're epistemically rational supervenes on facts that are 'internal' to you. The relevant sense of 'internal' can be cashed out in a variety of ways. If there are diachronic norms of epistemic rationality, then whether you're epistemically rational *now* is determined in part by your past epistemic states. And facts about the past are not, in the relevant sense, internal to you.

The proponent of diachronic norms faces a dilemma. We can't endorse both of the following claims: that epistemic rationality imposes cross-temporal constraints on belief, and that epistemic rationality is determined only by what is 'internal' to the agent.

Faced with a choice between diachronic norms and epistemic internalism, I will argue that we should choose diachronic norms. I argue that the rejection of diachronic norms incurs a number of serious problems: most notably, that it permits discarding evidence, and that it treats agents who are intuitively irrational as epistemic ideals.

Here is how the paper will proceed: in section 1, I'll explain the framework in which much of my discussion takes place, i.e., the Bayesian view of rationality. Then I'll introduce in more detail the objection to diachronic epistemic norms, some of its common motivations, and how the debate is situated within epistemology.

In section 2, I offer three objections to the synchronic-norms-only view. In 2.1, I argue that time-slice rationality entails that discarding evidence is rational. 2.2 argues that there are intuitive normative differences between agents who conform to diachronic norms and those who don't. The opponent of diachronic norms is committed to a strong claim: that no agent can ever be worse than another in virtue of purely diachronic differences between them. There are intuitive counterexamples to this generalization. In 2.3, I argue that according to an attractive view in philosophy of mind, all irrationality is fundamentally diachronic. So the synchronic-norms-only view may wind up committed to there being no epistemic rationality at all.

In section 3 I discuss the motivates, explicit or tacit, of the synchronic-norms-only view. I discuss the idea that cognitive limitations somehow limit our epistemic liability in 3.1. In 3.2 I discuss the idea of epistemic *ought*-implies-*can* and epistemic *responsible*-implies-*can*. 3.3 describes a notion of relative rationality, which allows us to accommodate many of the intuitions cited in favor of the synchronic-norms-only view.

Section 4 discusses an objection to diachronic norms prohibiting information loss. What if one can ensure a net gain in information only at the cost of losing some information? I discuss diachronic norms that can accommodate the idea that this sort of 'information trade-off' can be rational. I conclude briefly in section 5.

1 The conflict

1.1 Bayesianism

Before I begin, let me state some background assumptions. First, I will assume a partial belief framework. (Nothing hinges on this.) On this view, beliefs come in degrees (where a degree of belief is called a ‘credence’). Credences fall in the interval $[0, 1]$, where credence 1 represents certain belief, credence 0 represents certain disbelief, credence $\frac{1}{2}$ represents maximal uncertainty, and so on. A person’s total belief state is represented by a credence function, i.e. a function from propositions to real numbers in $[0, 1]$.

According to the classical Bayesian picture, there are two kinds of coherence that rational credences exhibit, one synchronic and one diachronic. The synchronic constraint is known as Probabilism:

Probabilism: Rational credences form a probability function: that is, they obey the following three axioms. Where \mathcal{W} is the set of all worlds under consideration¹:

1. *Nonnegativity:* for all propositions $A \subseteq \mathcal{W}$, $Cr(A) \geq 0$
2. *Normalization:* $Cr(\mathcal{W}) = 1$
3. *Finite additivity:* if A and B are disjoint, then $Cr(A \vee B) = Cr(A) + Cr(B)$

The diachronic constraint is known as Conditionalization:

Conditionalization: let E be the strongest proposition an agent learns between t and t' . Then the agent’s credences should update such that $Cr_{t'}(\cdot) = Cr_t(\cdot \mid E)$, where $Cr(A \mid B)$ is usually defined as follows:

$$Cr(A \mid B) = \frac{Cr(A \wedge B)}{Cr(B)}$$

Conditionalization has two basic effects: first, you treat all possibilities (that is, worlds) that are incompatible with your new evidence as dead. They are given credence 0. Second, you reappportion your credences among the remaining live possibilities, preserving relative proportions between the possibilities.

¹ Throughout I will be assuming that credence functions range over subsets of a finite set of worlds.

Now, one of the consequences of Conditionalization is that once you rationally learn something, you can't rationally unlearn it. You can't rationally lose information. (The set of live possibilities only shrinks.) This is, as stated, a strong and fairly controversial constraint.

There are analogs to Conditionalization in the full belief framework. For example, Jane Friedman (manuscript), defends the following norm of inquiry: when a question has been closed, don't reopen it. This is a close analog to Conditionalization's controversial consequence: that possibilities with credence 0 cannot recover positive probability. There are other diachronic norms that are weaker: for example, some forms of epistemic conservatism say that if you rationally believe a proposition at an earlier time, then it remains rational for you to continue believing it at later times, as long as you don't receive any new, disconfirming evidence.

I want to offer a general diachronic norm that cross-cuts whether we treat belief states with the full belief framework or the partial belief framework, and also cross-cuts whether we treat the overriding diachronic norm as Conditionalization, or whether we accept alternatives diachronic norms on credences (e.g. Jeffrey Conditionalization). Here is a candidate:

Diachronic evidentialism: An agent should only change her epistemic state by updating on new evidence.

Note that this is, on its face, a fairly strong norm. One needn't endorse this strong a norm in order to believe that there are diachronic constraints on rationality. But we'll start with something this strong, and see what can be said in favor of it.

First, though, we should consider objections to diachronic norms.

1.2 The rejection of diachronic rationality

Sarah Moss (2012) describes a 'general movement' towards rejecting diachronic norms of rationality. The aim of this movement: to take statements of diachronic norms like Conditionalization and replace them with analogous synchronic norms. According to Moss:

It is naïve to understand Conditionalization as a diachronic rule that says what credences you should have at a later time, given what credences you had at an earlier time, literally speaking. Instead we should understand it as a synchronic rule... Of course, one might claim that Conditionalization was originally intended as a literally diachronic rule, and that 'Conditionalization' should therefore be reserved for a

rule that binds together the credences of different temporal slices of agents—but I am inclined to interpret the Founding Fathers charitably. (Moss, 2012, 24)

Opponents of diachronic epistemic norms include Talbott (1991), Christensen (2000), Williamson (2000), Meacham (2010), and Hedden (2012).

There are a variety of motivations for a synchronic-norms-only epistemology. Some, e.g. Williamson, simply find diachronic constraints like Diachronic Evidentialism implausible. For others, the synchronic-norms-only view follows from a more general principle—in particular, some form of epistemic internalism. Here, for example, is Meacham (2010):

In Bayesian contexts, many people have appealed to implicitly internalist intuitions in order to support judgments about certain kinds of cases. But diachronic constraints on belief like conditionalization are in tension with internalism. Such constraints use the subjects beliefs at other times to place restrictions on what her current beliefs can be. But it seems that a subjects beliefs at other times are external to her current state. (87)²

There are a number of different forms of epistemic internalism. The two varieties that are perhaps most familiar are *mentalistic internalism* and *access internalism*.

Mentalistic Internalism: the facts in virtue of which a subject is epistemically rational or irrational supervene the subject's mental states.³

Access Internalism: the facts in virtue of which a subject is epistemically rational or irrational supervene on those of the subject's mental states that she's in a position to know she is in.

It's worth noting that neither of these immediately conflicts with diachronic constraints on rationality, at least as stated. After all, it might be that what's rational

² Note that while Meacham argues that there is a conflict between Conditionalization and internalism, and provides a synchronic alternative to Conditionalization, he is (at least in his (2010) not committed to the denial of traditional diachronic Conditionalization.

³ Note that this is (at least arguably) orthogonal to internalism about mental content. It's consistent to hold that whether an agent's beliefs are rational is determined by what's in the head, while at the same time holding that the correct characterization of the *contents* of an agent's beliefs will involve

for an agent believe at one time supervenes on her mental states at another time, or her mental states at many different times, or those mental states that she has access to at many different times, etc.

Opponents of diachronic norms often appeal to a form of access-internalism: facts about our past mental states are irrelevant to our current rationality because they are, at least in some circumstances, inaccessible to us.⁴ (A mental state is *accessible* to an agent iff, if the agent is in the mental state, then she is in a position to know that she is.) And so the internalist objection to diachronic rationality is best interpreted as involving the following form of internalism:

Time-Slice Internalism: the facts in virtue of which a subject is epistemically rational or irrational *at a particular time t* supervene on those of the subject's mental states that she's in a position to know she is in *at t* .

Here's an example statement of this sort of internalism:

Whether it is rational to retain or abandon a belief at a time is a matter of which of these makes sense in light of your current epistemic perspective, i.e., in light of what you currently have to work with in revising your beliefs. (McGrath, 2007, 5)

Time-slice internalism immediately entails that the norms governing epistemic rationality are purely synchronic.

The motivations for time-slice internalism draws on an analogy between the past and the external: our access to our past mental states is, at least in principle, limited in just the same way as our access to the external world.⁵ The fact that we had certain mental states in the past does not entail that we are, at present, in a position to know that we had those mental states.

We can show the differences between time-slice internalism and traditional access internalism by appeal to different forms of skeptical scenario:

Example #1

Suppose there are two agents who have exactly the same mental states.
Furthermore, both agents have access to exactly the same mental states.

appeal to the agent's environment.

⁴ Williamson is, of course, an exception, since he is not an internalist of any sort. Christensen's objection to diachronic norms, which I discuss in section 4, doesn't require appeal to any form of internalism.

⁵ Meacham (2010), Hedden (2012).

But one agent has mostly true beliefs about the external world; the other is a brain in a vat and is systematically deceived about the external world.

The internalist intuition about this case: if the undeceived agent is rational, so is the brain in the vat.

The time-slice internalist invites us to make the analogous judgment about an agent who is systematically deceived not about the external world, but about her past memories:

Example #2

Suppose there are two agents who have exactly the same mental states at a particular time t . Furthermore, both agents have access to exactly the same mental states. But one agent has mostly true beliefs about her past memories; the other has a brain implant that dramatically alters her beliefs, memories (or, if you like, quasi-memories), and other mental states erratically, and so at t she is systematically deceived about her past beliefs.

The question is: should these cases be treated as epistemically analogous? Do we have the same kind of intuition that, in the second example, if the ordinary agent is rational, then the memory-scrambled agent is rational? I would find it surprising if anyone claimed to have strong intuitions about whether the latter agent is rational.

The proponent of synchronic-norms-only rationality emphasizes the analogy between the agent who's deceived about the external world and the agent whose memories are regularly scrambled. After all, they are both doing the best they can under strange, externally imposed circumstances.

The proponent of diachronic norms responds that the scrambled agent should instead be understood on analogy to someone who is given a drug that makes him believe contradictions. They are both doing the best they can under strange, externally imposed circumstances—but nevertheless, they are not ideally rational. I'll argue for this claim in greater detail in section 3.

1.3 Orienting the debate

I'm concerned to defend a fairly weak claim: that there are diachronic norms of epistemic rationality. Advocating diachronic epistemic norms does not entail advocating Conditionalization, which is clearly an extremely strong constraint.

To orient the debate over diachronic norms, we can consider various kinds of loose (!) alliances. The debate is in some ways aligned in spirit with the debate over

epistemic externalism v. internalism, for obvious reasons: if there are genuinely diachronic epistemic norms, then whether a belief state is rational at a time can depend on facts that are inaccessible to the agent at that time.

There are also some similarities in spirit between defenders of diachronic norms and defenders of epistemic conservatism. According to epistemic conservatism (at least, of the traditional sort; there are, of course, varieties of conservatism), if you find that you have a belief, that provides some (defeasible) justification for continuing to have that belief. One way of drawing out this analogy: the epistemic conservatist holds that if an agent rationally believes that p at t , then it is (*ceteris paribus*) *permissible* for the agent to believe that p at a later t' .⁶ The defender of a diachronic norm like Conditionalization holds that if an agent rationally believes that p (with certainty) at t , then she is rationally *required* to believe that p at t' .

But it's worth noting that there are weaker diachronic requirements that could constrain rational belief: for example, that one shouldn't reduce or increase confidence in a proposition (in which her previous credence was rational) unless she receives new evidence *or* forgets evidence. The time-slice internalist is, therefore, endorsing a fairly strong claim. As I'll argue in the next section, there are costs to denying that rationality imposes *any* diachronic constraints on belief.

2 Problems for time-slice rationality

2.1 Problem #1: permissibly discarding evidence

One of the benefits that time-slice internalists claim for their view is that, by rejecting Conditionalization, they are able to vindicate the idea that forgetting doesn't make a person irrational. If Conditionalization applies, without qualification, over the whole of an agent's life, then any instance of forgetting would be sufficient to make the agent irrational.

The flip side is that time-slice internalism also makes any instance of discarding evidence epistemically permissible. And discarding evidence is a canonical example of a violation of epistemic norms. The reason that time-slice internalism has this effect is that discarding evidence is a fundamentally diachronic phenomenon. At some time, you receive evidence. At a later time, your attitudes fail to reflect the fact that you've received that evidence.

Example #3

⁶ See e.g. (Burge, 1997).

Suppose an agent has strong beliefs about whether capital punishment has a deterrent effect on crime. Then he learns of a study that provides evidence against his view. So he should reduce his confidence in his belief. But instead our agent (involuntarily) discards the evidence; he loses any beliefs about the study; it has no enduring effect on his attitudes regarding capital punishment. Now he can go on confidently endorsing his beliefs without worrying about the countervailing evidence.

This is a standard example of irrationality. (One might object: an agent like this is epistemically irrational only if he voluntarily discards the evidence. But cognitive biases are not voluntary; so this objection would have the consequence that cognitive biases never result in irrational belief. I take this to be uncontroversially false.)

Discarding evidence is epistemically irrational. Therefore there are diachronic norms of epistemic rationality. There's not much more to say about this. But to my mind it is a serious challenge to the synchronic-norms-only view; perhaps the most serious.

2.2 Problem #2: deviating from epistemic ideals

Some kinds of belief change are plausibly described as deviating from some sort of epistemic ideal, even when no synchronic norms are violated. It might be controversial whether, by virtue of deviating from the ideal, the agent is irrational. But given that there are purely diachronic epistemic ideals to deviate from, it follows that there are diachronic epistemic norms.

Consider again an agent whose total belief state is entirely overhauled at regular, and perhaps frequent, intervals (every minute? every second?). At every instant her credences are probabilistically coherent. And they uphold any other synchronic constraints on rational belief: for example, they are appropriately sensitive to chance information; they reflect whatever the epistemically appropriate response is to whatever phenomenological inputs the agent has at that instant; etc. However strong you make the norms of synchronic rationality, our agent obeys all of those norms at each instant.

But her total belief state at one moment is largely different from her total belief state at the next. If you asked her a minute ago where she was from, she'd say Orlando; if you asked her now, she'd say Paris; if you ask her a minute from now, she'll say Guelph. These changes are random.

The time-slice internalist is committed to the claim that our agent is *ideally rational*. I think this is false. Whether or not the agent rises to the level of rationality, it is clear that she is epistemically sub-ideal: she is doing worse, epistemically, than someone whose credences are more stable over time.⁷

Objection: If her evidence changes with each belief overhaul, then perhaps it is rational for her to overhaul her beliefs so frequently.

Reply: In order to assess whether her evidence changes with each belief overhaul, we would need to say more about what ‘her evidence’ is. For example, if you believe her evidence is what she knows,⁸ then no doubt it’ll overhaul, since her beliefs overhaul. That doesn’t get us any further toward defending the claim that she is rational. It might just be that she irrationally stops believing various propositions that she previously knew.

If you believe her evidence is something else—perhaps something to do with phenomenology—then in order to ensure that the example is one where she obeys synchronic norms, this will have to overhaul regularly too. But let’s note that phenomenology a very, very thin source of information. If you think that memorial phenomenology is the basis for your beliefs—consider how few of those beliefs can be justified on the basis of memorial phenomenology at any given moment. There is a limit to how much phenomenology you can conjure up in a short time span. To the extent that I understand this sort of view, it seems to me that it is susceptible to the same charge that the time-slice internalist presses against the defender of Conditionalization: that it declares us all irrational.

On the other hand, if diachronic evidentialism is correct, then ‘her evidence’ is all the evidence she has received, not just the evidence that is accessible to her in the moment. The diachronic will say: her evidence does not dramatically change, and therefore it’s irrational for her beliefs to dramatically change.

Now, one can agree with me that the agent with erratically shifting beliefs is epistemically non-ideal, and still judge the agent to be rational. One might, for example, have a *satisficing* view of rationality: maybe it isn’t necessary to perfectly satisfy all epistemic norms in order to be epistemically rational. This kind of view isn’t common among Bayesians, who tend to accept that rationality just is ideal rationality, and who tend to accept happily that none of us is rational. But I take it that this is a common presupposition in informal epistemology. For example,

⁷ Of course, it’s entirely appropriate that an agent’s beliefs should continuously change *a little* all the time: she should update on new information about, e.g., the passage of time, new events that she encounters, etc. But in the example I’m concerned with, a much greater proportion of her beliefs change, and not simply because she’s exposed to new evidence.

⁸ See (Williamson, 2000) for the canonical defense of this identity.

informal epistemologists typically accept that it's not rationally required that we believe all the consequences of our beliefs, though we would be rationally better if we did.

As long as we accept that the agent whose credences shift erratically is doing worse, epistemically speaking, than the agent whose credences only change by rational update on evidence, then for my purposes, it doesn't matter whether we call the former agent irrational or rational. We have already admitted that there are diachronic epistemic norms.

2.3 Problem #3: all incoherence is diachronic

Some form of functionalism about belief is widely held among formal epistemologists. Probably the most influential form is **interpretivism**, the view that your belief state is simply the body of credences that best rationalize your behavioral dispositions.

One effect of this sort of functionalism is that the kinds of facts in virtue of which you believe A at a particular time t are also facts in virtue of which you don't believe $\neg A$ at t . Similarly, the facts in virtue of which you have credence .6 in B at t also make it the case that you have credence .4 in $\neg B$ at t . The facts that make it the case that you've received evidence that C at t are facts that make it also the case that you increase your credence in C at t . And so on.

How, then, can we ever correctly attribute incoherence to anyone? Agents can be 'fragmented': they can, in effect, believe different things relative to different contexts, or for different purposes. For example, an agent may exhibit a belief that A in her linguistic behavior in some contexts, and yet manifest a belief that $\neg A$ in her non-linguistic actions in another context.

One needn't accept the full fragmentationist package in order to accept that the very facts that make it the case that a person believes A at a particular time also make it the case that he rejects $\neg A$ at that time; and similarly for other sorts of relations that beliefs stand in.

An effect of this view: in a particular context, at a particular time, an agent is *always* synchronically coherent. Synchronic coherence, on this interpretation, is either a trivial norm, or else a norm constraining belief-attribution rather than belief itself.

If this view in philosophy of mind is correct, opponents of diachronic rationality are pushed in to a corner. They must either reject this attractive philosophy of mind, or else reject the idea that there are *any* substantive epistemic constraints on belief.

Neither of these is an attractive option.⁹

3 Epistemic ‘blamelessness’ does not entail epistemic ideality

3.1 Diachronic evidentialism and information loss

Let me emphasize again: I am concerned primarily with defending Diachronic Rationality, the claim that there are diachronic epistemic norms. Here are two stronger claims:

Rationality = Ideal Rationality In order to be epistemically rational, one must perfectly satisfy all epistemic norms, synchronic or diachronic;

or even stronger:

Rationality Requires Lifelong Conditionalization In order to be epistemically rational, you must satisfy Conditionalization over the entire course of your life.

There are a variety of ways we could resist these extensions of Diachronic Rationality. For example, you might accept that satisfying Conditionalization would make an agent epistemically better but that it isn’t always necessary for rationality; perhaps there are sometimes extenuating circumstances. Or you might accept that Conditionalization is rationally required over stretches of time, but not an agent’s entire life. (Perhaps it’s required between instances of some psychological event of forgetting, where this might be psychologically distinguished from discarding evidence.)

The real question is: who is making the universal claim and who is making the existential claim? The opponent of diachronic norms insists that no one is *ever* irrational by virtue of diachronic facts. All we need to convince ourselves that this view is false is one instance where, e.g., discarding evidence is epistemically sub-ideal.

⁹ *Objection:* fragmentationists might also hold that individual fragments also necessarily update by Conditionalization, and so for each fragment this norm is also trivial. *Reply:* One can exhibit irrationality diachronically in other ways than by failing to conditionalize *within* a fragment. For example, fragments can conflict with each other in such a way as to make the agent irrational. But the facts in virtue of which this is true have to do with different fragments characterizing the agent’s belief state manifest at different times.

Nevertheless, I want to explore a defense of diachronic evidentialism, the comparatively strong claim that epistemically, we should only change our beliefs by updating on new evidence. We'll set aside the question of whether an agent who violates diachronic evidentialism is irrational in all circumstances.

One form of diachronic evidentialism is Conditionalization. A common complaint against Conditionalization is that it entails that forgetting something learned with certainty is irrational. This result is often met with an incredulous stare; counterargument is treated as unnecessary.

Forgetting is not irrational; it is just unfortunate. (Williamson, 2000, 219).

It seems to me that forgetting is not just unfortunate but *epistemically* unfortunate. And 'epistemic misfortune' is simply a gentler name for epistemic sub-ideality. In any case, even if Williamson is correct, it may still be that Conditionalization has epistemic normative force.¹⁰

Let me acknowledge: I'm not concerned about whether we accept the claim I above called 'Rationality = Ideal Rationality.' Where we draw the line between epistemic trespasses that are sufficient for irrationality and those that aren't doesn't seem to me obviously substantive. Sociologically speaking, formal and informal epistemologists tend to talk about rationality in quite different ways. For many informal epistemologists, to be 'irrational' is to be (at least a little) insane; the majority of us are by and large rational. It is common to think, e.g., that one is not rationally required to believe all the consequences of one's beliefs (even though perhaps by doing so you'd be epistemically better). By contrast, among formal epistemologists, it is more common to use 'irrational' to mean *rationality imperfect*. To be epistemically 'irrational', in their sense, is to deviate from epistemic ideals.

Now, whether or not we call it 'irrational', forgetting—losing information—deviates from our epistemic ideals. Compare it with other epistemic ideals:

Deductive closure: if an agent believes A and $A \vdash B$, then the agent should believe B .

¹⁰ There are complaints against Conditionalization that have nothing to do with forgetting: for example, that it only allows update when our evidence provides us with credence 1 in some new proposition (unlike, e.g., Jeffrey Conditionalization), and that it doesn't allow us to lower our credence in any proposition from 1 even in circumstances where no forgetting takes place (e.g. in Artzenius's (2003) Shangri-La example). But neither of these objections extends to Diachronic Evidentialism; so these considerations simply invite us to find a suitable diachronic replacement for Conditionalization.

Because of our cognitive limitations—for example, the fact that we can't believe all mathematical truths—actual agents' beliefs are never actually closed under deduction.

Probabilism creates similar problems:

Probabilism. Our credences must form a probability function.

This entails that we must, e.g., have credence 1 in all necessary truths. It also entails that we must have infinitely precise credences: that there be a difference between having credence .2 and credence .20000000000001. But because of our cognitive limitations (we are finite beings!), actual agents never actually have infinitely precise credences.

It should be a familiar point, then, that because of our cognitive limitations, no actual agents are epistemically ideal. And there's no obvious reason to treat forgetting any differently. Actual agents' forgetfulness is just another cognitive limitation that stands in the way of epistemic ideality.

3.2 Epistemic *ought* implies *can*?

Now, one might object: so much the worse for any of these norms! Surely we're not blameworthy for beliefs that result from our cognitive limitations. If you *can't* satisfy the norm, then the norm doesn't apply to you. (After all, *ought* implies *can*.)

But this is simply false. Our friend in his tinfoil hat can't make himself stop overtly believing contradictions. That doesn't make him *epistemically ideal*. It is a commonplace in epistemology that sometimes a person can be irrational even when he is 'doing the best he can'.

Even if the epistemic *ought*-implies-*can* argument were successful against ideals like deductive closure, probabilism, or precise credences, it's not clear how it is supposed to apply to forgetting. After all, a norm against forgetting would say: if you're in a certain kind of state, you should *continue* to be in that state. In the other cases, no actual agents can be in the relevant state in the first place. So it's not as though it's psychologically or physically impossible for you to be in the recommended belief state. It's just that you can't always *make yourself* remember something.

But in epistemology *ought* ϕ doesn't imply *can make yourself* ϕ . It's not as though you can simply make yourself believe anything. (Try believing that I am a goat!) Beliefs are not under our immediate voluntary control.¹¹ And so if there's

¹¹ We can take actions to induce beliefs, e.g. gathering evidence, or take actions to slowly indoctrinate

any sense in which *ought*-implies-*can* in epistemology—which is doubtful—it does not apply in the case of forgetting.

This point generalizes to any argument for time-slice internalism that appeals to the idea that we cannot be responsible or blameworthy for believing in accordance with past evidence that we no longer have immediate access to. Epistemic rationality has nothing to do with responsibility or blameworthiness.

Perhaps the greatest challenge for the time-slice internalist is to justify their view in some way that doesn't appeal to some misguided epistemic *ought*-implies-*can* or epistemic *ought*-implies-*responsible* principle.¹²

3.3 Relative rationality

One fear we might have about accepting epistemic principles that ordinary agents can't perfectly realize is that we would then have to accept that the norms of rationality are, in some sense, only for ideal agents; they don't apply to any actual agents.

But that's rather like saying that if you're not ideally law-abiding—you've already gotten a speeding ticket; there's nothing you can do to change that fact—then traffic laws no longer apply to you. Suppose the traffic laws say:

1. Don't get speeding tickets;
2. If you get speeding tickets, pay the speeding tickets;
3. If you don't pay your speeding tickets, go to your court hearing;
4. ...

Then this set of legal norms generates different 'levels' of law-abidingness. 'Ideal law-abidingness' amounts to obeying *all* of these (where everything after 1 you satisfy trivially by virtue of satisfying 1). Still, if you *can't* obey all of the laws, you're legally required to obey 2, 3, ...; and if you *can't* obey 2, then you're legally required to obey 3, etc.. What the traffic laws require of you in particular circumstances is relativized to what you are capable of. Still, though, if you are not capable of satisfying all of the laws, then you are not *ideally* law-abiding.

We can represent the norms of rationality as having a similar structure:

ourselves over time. But generally speaking, we cannot believe something merely by trying.

¹² In a 2012 AAP talk (no manuscript currently exists), Wolfgang Schwarz argued, similarly, that the motivation for rejecting diachronic norms derives from the idea that they cannot be action-guiding, and this turns on an illicit conflation of the practical with the epistemic.

1. Be diachronically *and* synchronically coherent.
2. If you can't be both, be synchronically coherent.
3. ...etc.

So, like law-abidingness, we can think of rationality as *relative*—in particular, relative to our cognitive limitations. Ideal rationality is a special case of relative rationality: it is the case where there are no limitations.

3.4 Rationality vs. epistemic ideality?

I have emphasized that there's a clear sense in which the subject who violates diachronic norms is doing worse, epistemically, than the subject who doesn't. But the time-slice internalist might object: the person who happens to *know* less is also doing worse, epistemically, than a person who knows more. But that doesn't mean that the person who knows less is *irrational*. So, the time-slice internalist might conclude, not all epistemic norms are norms of rationality.

There is a natural way of drawing a distinction between norms of epistemic rationality vs. other epistemic norms. In the practical realm we sometimes distinguish 'objective' and 'subjective' norms. By analogy, we might consider it an objective epistemic norm that we should believe all true propositions and disbelieve all false propositions, in the same way that according to the utilitarian, it is an objective norm that we should maximize utility. And conversely, we might think of norms of rationality as including only the subjective norms. Where do diachronic norms fall on this divide? Which of the epistemic norms are norms of epistemic rationality?

As a working hypothesis, here is my suggestion: we should think of the norms of epistemic rationality as those that characterize the states of the *agent* and not her environment. One of the ways of cashing this out: the epistemic norms are the constraints that characterize the epistemic states of the ideal information gatherer.¹³ The ideal information gatherer is non-omniscient; none of her beliefs is guaranteed to be true except on the basis of evidence.¹⁴

¹³ Schwarz defended Conditionalization with this analogy: suppose we want to build a robot to gather information for us in whatever environment he ends up in. We have the option of programming it to obey diachronic evidentialism. Should we? It seems fairly obvious that we should: then the robot will not lose information, and so will end up with more information.

¹⁴ This is, again, a working hypothesis. But there is another answer that I'm sympathetic to. There's a few suggested in semantics, but separable from linguistic considerations, that there is no real sub-

Epistemic rationality involves having beliefs that approximate the truth as much as possible, given our non-omniscience. On this view, though, there's no reason to think of diachronic norms as somehow external to rationality. Retaining information will, by and large, help you keep your belief state more accurate.

4 Rational information loss

4.1 Losing information to gain information

Now, it can't be that losing information *necessarily* makes your beliefs less accurate. For example: suppose that, by chance, you happen to forget only misleading evidence. Then losing information actually makes your beliefs more accurate. Rather, retaining information makes it more likely that your credences will be more accurate, roughly speaking. It increases the *expected accuracy* of your credences. (I will say more about this in section 1.)

Now, conditionalizing on new information is an example of pure information gain. And forgetting and discarding evidence are examples of pure information loss. But what should we say about mixed cases?

We can define an **information trade-off** as a case where you gain some information at the cost of losing some other information. If taking an information trade-off can be rational, then some forms of diachronic norm are false. For example, Conditionalization is false: rational informational trade-offs would require rational information loss. Christensen (2000) uses an example with the following structure to argue against the view that there are diachronic epistemic norms:

Example #4

Suppose you know that someone knows more than you about some topic. You know some things she doesn't know, but on the whole she's more informed on the topic. It would be gauche to ask her about the topic. Luckily, you have the option of using a credence downloader to replace your credences on the topic with hers. Is it permissible for you to do so?

Christensen invites us to judge that it is indeed permissible.

jective/objective divide. There are simply grades of more or less subjective norm. Likewise, it might be that there's ultimately no defensible line to be drawn between subjective and objective epistemic norms. Diachronic norms are more 'subjective' than the norm of truth, but more 'objective' than some synchronic norms (like probabilism), which are in turn more 'objective' than other synchronic

Now, it should be clear that this is at best an argument against *some* diachronic norms, not against diachronic rationality in general. But one interesting fact about this case is that if you take the trade-off, you violate Conditionalization—but you also increase the expected accuracy of your credences. So, if epistemic rationality consists in maximizing expected accuracy, then Conditionalization can't be a norm of epistemic rationality.

Now, there are two possible objections one could make against Conditionalization on the basis of an example like this.

Objection #1. Taking the trade-off maximizes expected accuracy, so you're rationally required to violate Conditionalization.

This shouldn't trouble the proponent of Conditionalization. The norms of epistemic rationality govern only epistemic states, not actions like using a credence downloader. If we were rationally required to perform actions that maximize the expected accuracy of our credal states, then we would, for example, be rationally required to perform constant experiments, to read all of Wikipedia, etc.

Objection #2. If you do take the trade-off, your resulting epistemic state is rational. So it must be permissible to violate Conditionalization.

This objection is more troubling for the proponent of Conditionalization. If this objection is correct, then Conditionalization is false. At most Conditionalization holds across periods of time where no informational trade-offs are available.

There are the two options, then, for the proponent of diachronic norms:

1. We can stick with Conditionalization and reject the claim that there are epistemically rational informational trade-offs. (We might concede that informational trade-offs are still pragmatically rational.)
2. Alternatively, we can adopt diachronic norms that are more liberal than Conditionalization.

There's something to be said for both options and I won't defend one over the other. There's little more to be said about the first option, though, so let's explore the second option. But first, we should say a little bit more about what expected accuracy is.

norms, and so on.

4.2 Epistemic utility theory

Epistemic utility theory formalizes the notions of the accuracy and the expected accuracy of a credence function. The aim of epistemic utility theory is to use the tools of decision theory, combined with an epistemic version of value, in order to give a foundational justification for various epistemic norms.

The most widely discussed epistemic utility functions are *gradational accuracy* measures. The accuracy of a credence is its nearness to the truth (by some measure). A credence function with maximal accuracy would assign credence 1 in all truths and credence 0 in all falsehoods. In other words, it would be omniscient.

Decision rules are adapted from decision theory, e.g. expected utility maximization. Paired with accuracy as the relevant measure of utility, we end up with the decision rule:

Maximize Expected Accuracy: adopt the credence function that has the highest expected accuracy, by your own lights.¹⁵

The expected accuracy of a credence function is standardly calculated as the sum of a credence function's accuracy in each world, weighted by the probability of that world. In symbols:

$$EU^{Cr}(Cr') = \sum_{w_i \in \mathcal{W}} Cr(w_i)U(Cr', w_i)$$

With the decision rule Maximize Expected Accuracy, various results can be proven. Call a function from evidence to credence functions an 'update policy.' Greaves & Wallace (2006) and Leitgeb & Pettigrew (2010b) proved that from an agent's own perspective, given the choice of all possible update policies, Conditionalization alone maximizes expected accuracy. So, one might conclude hastily, in order to be an ideal information gatherer, your credences should update by Conditionalization.

But, you might ask, isn't example #4 intuitively a case where I know that some other credences than my own maximize expected accuracy from my point of view? In that example, I would receive an increase in expected accuracy if I updated by some means that would involve violating Conditionalization. Does that example conflict with the results of epistemic utility theory?

¹⁵ In (Carr, manuscripta), I call into question the idea of expected accuracy used by epistemic utility theorists. For the purposes of addressing this objection to diachronic rationality, though, I will take the appeal to expected accuracy at face value.

4.3 Assessing rational trade-offs

In fact, there's no conflict between the idea that there could be rational information trade-offs (violating Conditionalization) and the epistemic utility theoretic result that Conditionalization is the only update policy that maximizes expected utility.

The reason: update policies are narrowly defined as functions from evidence to particular credence functions. But it's a feature of information trade-offs that you do not know, in advance, what credences you will adopt as a result of taking the trade-off. (If you did, then you could update on that information directly, which would then amount to pure information gain.) Indeed, on common assumptions,¹⁶ it cannot be the case for any particular credence function that you can rationally assign it higher expected accuracy than your own credence function. But if you have the option of adopting *whichever* of a set of possible credence functions that has updated on some information (information that is not otherwise accessible to you), then that option can maximize expected accuracy from your perspective.

Let's consider a particular case of an informational trade-off, specifying some of the details from example #4. Suppose a particular coin is either fair or biased (with a $\frac{3}{4}$ heads bias), and it will land either heads or tails. You are uncertain about both matters. Now, you and your colleague start with the same priors:

$$\begin{array}{lll} w_{FH}: \text{fair, heads} & Cr_0(w_{FH}) & = \frac{1}{4} \\ w_{FT}: \text{fair, tails} & Cr_0(w_{FT}) & = \frac{1}{4} \\ w_{BH}: \text{biased, heads} & Cr_0(w_{BH}) & = \frac{3}{8} \\ w_{BT}: \text{biased, tails} & Cr_0(w_{BT}) & = \frac{1}{8} \end{array}$$

Then you learn whether the coin lands heads or tails. Your colleague learns whether the coin is fair or biased. Both of you conditionalize on your respective evidence. You are not permitted to know the answers to both questions.

Suppose you learn that the coin lands heads. You have a credence downloader that will allow you to perform the informational trade-off. Is it epistemically rational for you to give up your knowledge in order to gain your colleague's?

Applying the rule Maximize Expected Accuracy isn't straightforward. Since we don't know what your colleague has learned, we don't know which credence function to assess. So it's not obvious how we can even determine the expected accuracy of your colleague's credence function.

¹⁶ Namely, that epistemic utility functions must be *proper* in the sense that they yield the result that any coherent credence function maximizes expected accuracy by its own lights.

Here is my suggestion: we can introduce a new kind of epistemic action. Call it *learning the answer to a question*. Learning the answer to a question involves taking an epistemic option when you're not in a position to know what credence function it will result in your adopting.¹⁷

For a question \mathcal{Q} (i.e. a partition over the set of epistemically possible worlds), let $Cr_{\mathcal{Q}}$ be Cr_0 conditionalized on whatever the true answer to \mathcal{Q} is (that is, whichever proposition in \mathcal{Q} is true at the world of assessment).

In our example, we can call whatever credence function your colleague has after learning whether the coin is biased or fair $Cr_{\mathcal{Q}_{BF}}$. Note that ' $Cr_{\mathcal{Q}_{BF}}$ ' is a description: it picks out different credence functions in different worlds. Ex hypothesi, your colleague updates on B in B -worlds and on F in F -worlds.

Now, with a concrete example in hand, and a new tool (the epistemic act of learning the answer to a question), we can ask: should you take the trade-off? We need to explain how to calculate the expected accuracy of $Cr_{\mathcal{Q}_{BF}}$ from your point of view:

1. Calculate the accuracy of Cr_B at B -worlds and Cr_F at F -worlds.
2. Sum the values, weighted by their probability according to Cr_H .

In symbols:

$$EU^{Cr_H}(Cr_{\mathcal{Q}_{BF}}) = \sum_{w_i \in \mathcal{W}} Cr_H(w_i)U(Cr_{\mathcal{Q}_{BF}}, w_i)$$

In this case, with plausible assumptions about the accuracy function U , taking the trade-off maximizes expected accuracy. Retaining your current credences does not.¹⁸

¹⁷ This kind of epistemic tool isn't just for science fictional cases where you are offered information trade-offs. We can do other things with our new epistemic acts. For example, they can be useful in decisions over whether it would be more informative to perform one experiment or another, in circumstances where it is impossible, or at least costly, to perform both.

¹⁸ Suppose U is the negative Brier Score. (Joyce (2009) and Leitgeb & Pettigrew (2010b) argue that only scoring rules like the Brier Score satisfy plausible constraints on epistemic utility functions.) Where $v_w(X) = 1$ if X is true at w and $v_w(X) = 0$ if X is false at w , $U(p, w) = -\sum_{X \subseteq \mathcal{W}} |v_w(X) - p(X)|^2$.

$$\begin{aligned} EU^{Cr_H}(Cr_{\mathcal{Q}_{BF}}) &= \sum_{w_i \in \mathcal{W}} Cr_H(w_i)U(Cr_{\mathcal{Q}_{BF}}, w_i) \\ &= Cr_H(w_{BH})U(Cr_B, w_{BH}) + Cr_H(w_{HF})U(Cr_F, w_{FH}) \\ &= -\frac{11}{40} \end{aligned}$$

This isn't surprising. Knowing that the coin landed heads isn't particularly informative about whether the coin is fair or biased, since it would be unsurprising either way. On the other hand, if you had instead learned that the coin had landed tails, then it would maximize expected accuracy to reject the trade-off. After all, knowing that the coin landed tails gives you fairly strong evidence in support of the coin's being fair.

So, we have a concrete case where taking an informational trade-off maximizes expected accuracy.

4.4 Discussion

As I said before, the defender of diachronic norms had two options for responding to an objection like this. If she continues to endorse Conditionalization, then she must reject the claim that it's rational to accept informational trade-offs. (This might involve rejecting the idea that we should perform those epistemic acts that maximize expected accuracy, or it might involve rejecting the idea that taking an information trade-off is an appropriately understood as an epistemic act.)

On the other hand, if we allow informational trade-offs as epistemic options, then accepting trade-offs can lead to maximizing expected accuracy. And if we accept that this is rational, then we can replace Conditionalization with amore liberal diachronic rule.

These two options provide us with different pictures of what an ideally rational agent's credences will look like over time. On the Conditionalization picture, the ideal rational agent's stock of information will only ever increase. But if we allow for violations of Conditionalization in informational trade-offs, then the ideally rational agent will in some circumstances take epistemic risks. These risks have two salient features that distinguish them from obeying Conditionalization. First, they involve sure loss of information; second, they may lead to decreases the agent's expected accuracy (from the perspective of her updated credences).

$$\begin{aligned}
 EU^{Cr_H}(Cr_H) &= \sum_{w_i \in \mathcal{W}} Cr_H(w_i)U(Cr_H, w_i) \\
 &= Cr_H(w_{BH})U(Cr_H, w_{BH}) + Cr_H(w_{HF})U(Cr_H, w_{FH}) \\
 &= -\frac{12}{25}
 \end{aligned}$$

The expected accuracy of your colleague's credence function ($Cr_{\mathcal{Q}_{BF}}$) is greater than the expected accuracy of your own credence (Cr_H). So if you know H , expected accuracy maximization requires you to take the trade-off.

Here is a candidate liberal diachronic norm (which is a variant on diachronic evidentialism):

Liberal norm: An ideally rational agent's credences change only in order to maximize their expected accuracy.

Note that for cases of pure information gain, Conditionalization will still hold. Furthermore, rational trade-offs arguably only occur in sci-fi cases.¹⁹ So, in ordinary cases, a more traditional, strict norm will still hold.

Strict norm: An ideally rational agent's credences only change in response to new evidence.

5 Conclusion

I've argued that there is a conflict between diachronic norms of epistemic rationality and a form of epistemic internalism. I've also argued that diachronically coherent agents are epistemically better. We should think of epistemic rationality as providing constraints that allow us to be more informed about our environment, whatever our environment happens to be like.

The diachronic norms I've advocated are at a middle ground between epistemic internalism and externalism: they are sensitive to facts that are external to the time-slice, but not necessarily external to the person. Contrast this sort of view with process reliabilism, which is concerned with whether some belief-forming process *actually* conduces toward the truth. Whether it does will depend on contingent facts about the agent's environment. A norm like expected accuracy maximization is concerned with whether an update method is *likely* to conduce toward the truth, by the believer's own lights.

If we take the option of maintaining Conditionalization, we are also given at a middle ground between epistemic conservatism and evidentialism. Like conservatism, Conditionalization permits us to continue to believe a proposition if we already believe it (with certainty). In fact, Conditionalization requires it. But unlike conservatism, Conditionalization doesn't permit continuing to believe a proposition after the evidence for it has been forgotten. Conditionalization requires remembering the evidence as well. In short, Conditionalization doesn't permit violations of diachronic evidentialism. Hence, what we're required to believe is always deter-

¹⁹ One might make the case that clutter avoidance is a more psychologically realistic version of an informational trade-off; see (Harman, 1986).

mined by what our evidence supports. It's just that our evidence—what we've learned—might escape us.

Chapter 3. What to Expect When You're Expecting

Epistemic utility theorists have argued that various epistemic norms can be given a decision-theoretic justification. By combining the tools of decision theory with an epistemic form of utility—accuracy—we can justify norms like conditionalization, probabilism, the principal principle, and others. These arguments generally use to one of two tools: the notion of *expected accuracy* and the notion of *accuracy dominance*. For example: it's been argued that the reason why we should have probabilistically coherent degrees of belief (“credences”) is that it's the only way to avoid credences that are accuracy dominated (Joyce, 1998, 2009), and the reason why we should update by conditionalization is that doing so uniquely maximizes expected accuracy (Greaves & Wallace, 2006; Leitgeb & Pettigrew, 2010b).

I'll show that deriving these results requires using notions of “dominance” and “expected utility” that are different in important respects from the dominance and expected utility used in standard (practical) decision theory. If we use the more familiar forms of expected utility and dominance, we can't justify the epistemic norms that epistemic utility theory had hoped to justify. Indeed, the prescriptions of epistemic utility theory can conflict with these norms.

I'm going to argue that the things epistemic utility theorists call “expected accuracy” and “accuracy dominance” can't really be the expected accuracy or accuracy dominance of epistemic acts in any conventional sense. It's not clear what they are; so far we don't have a good philosophical interpretation of these pieces of math. Without a philosophical interpretation, they are ill-equipped to do the epistemological work they were meant to do. For example, just telling us that conditionalization maximizes *this thing*—whatever it is—doesn't explain why we should conditionalize on our evidence. And so there are important open questions about the foundations of epistemic utility theory.

In short, those of us who are attracted to the project of epistemic utility theory face a dilemma. We must choose between old and new versions of norms like

Dominance and expected utility maximization. If we choose the old, decision-theoretic versions, then we vindicate the idea that rational belief has the aim of accuracy—but at the cost of giving up probabilism, conditionalization, and other plausible epistemic norms. On the other hand, if we choose the new norms, we can avoid predicting rational violations of probabilism, conditionalization, and so on—but at the cost of giving up the idea that epistemic rationality is a matter of having those credences that best approximate the truth, and thereby giving up the explanation for why we should obey these decision theoretic norms. Faced with this choice, I argue that we should choose the second horn. But this means that before epistemic utility theory can claim to justify epistemic norms like conditionalization or probabilism, it needs to explain and justify the normative claims it treats as fundamental. Otherwise, these “norms” are just uninterpreted formalism.

The plan for the paper is as follows: in section 1, I’ll briefly introduce epistemic utility theory and its motivations. Section 2 introduces the central example of the paper. The kinds of examples I’m interested in are cases where the truth of a proposition is not causally or evidentially independent of the credences you adopt.

In section 3, I’ll argue that with traditional decision-theoretic versions of expected utility, the norm that tells us to maximize expected epistemic utility (understood as accuracy) actually conflicts with conditionalization. How, then, were epistemic utility theorists able to argue that conditionalization always uniquely maximizes expected accuracy? The answer: the quantity that epistemic utility theorists have called “expected accuracy” is different in important ways from the measures of expected utility that we see in decision theory. I call this new measure “observational expected accuracy” or “OEA.”

In section 4, I show that with the traditional decision-theoretic conceptions of expected accuracy, maximizing expected accuracy also conflicts with probabilism.

In section 5, I argue that while OEA can generate better results than more familiar forms of expected accuracy (in that it preserves conditionalization and probabilism), it has a cost: it doesn’t have the same justification as other forms of expected accuracy. Indeed, I argue, it isn’t a measure of the expected accuracy of credal acts at all.

Section 6 turns to accuracy dominance arguments. I argue that closely related problems afflict them: what epistemic utility theorists have called the “dominance” relation is very different from the dominance relation in practical decision theory. Call this new relation “O-dominance.” While the prohibition of O-dominated credences doesn’t lead to conflicts with conditionalization or probabilism, it has a cost: it gives up the intuitive explanation we had for why dominated credences are epistemically bad.

In section 7, I consider the bigger picture. The reason why the familiar decision theoretic notions of expected utility—causal and evidential expected utility—and dominance generate conflicts with conditionalization and probabilism is that they conflate the epistemological ideal of respecting the evidence with the ideal of getting your credences close to the truth. I argue that a good philosophical interpretation of observational MaxExAcc and O-Dominance would be one that made explicit the relation between OEA-maximal or O-dominant belief and evidential support. If we want to answer the foundational questions about epistemic utility theory, this is where we should start.

1 Epistemic utility theory

A common way of modeling belief states is to treat beliefs as coming in degrees. Instead of a tripartite division of belief, disbelief, and suspension of judgment, we should represent belief states with degrees of belief (“credences”) in the interval from 0 to 1, where 1 represents certain belief and 0 represents certain disbelief. Total belief states are formally represented as functions from propositions (sets of possible worlds) to real values in the $[0, 1]$ interval. Credences are typically held to be regulated by two epistemic norms: probabilism and conditionalization.

Probabilism: A rational agent’s credences form a probability function: that is, they obey the following three axioms. Where \mathcal{W} is the set of all worlds under consideration¹:

1. *Nonnegativity*: for all propositions $A \subseteq \mathcal{W}$, $Cr(A) \geq 0$
2. *Normalization*: $Cr(\mathcal{W}) = 1$
3. *Finite additivity*: if A and B are disjoint, then $Cr(A \vee B) = Cr(A) + Cr(B)$

Conditionalization: let E be the strongest proposition² an agent learns between t and t' . Then a rational agent’s credences update such that $Cr_{t'}(\cdot) = Cr_t(\cdot | E)$, where $Cr(A | B)$ is standardly defined as $\frac{Cr(A \wedge B)}{Cr(B)}$.

The conjunction of these two claims is called “Bayesianism.”

The most familiar justifications for these norms are Dutch book arguments, which purport to show that agents with credences that violate either probabilism or

¹ Throughout I will be assuming that credence functions range over subsets of a finite set of worlds.

² I.e. the proposition such that all other propositions learned in the relevant span of time are entailed

conditionalization are susceptible to foreseeable exploitation. (For example, they'll accept sets of bets that jointly guarantee a sure loss.) But many have thought that this sort of justification is inadequate. We have good epistemic reason to have coherent credences, but Dutch book arguments give only pragmatic reasons for coherence.

Epistemic utility theory aims to provide non-pragmatic justifications for epistemic norms including probabilism and conditionalization. The basic tools of decision theory are given an epistemic flavor: the relevant sorts of "acts" are epistemic, i.e. adopting credence functions,³ and the relevant sort of value measured by the utility function is *epistemic* value. What do epistemic utility functions look like? The natural candidates are those utility functions that tie credences to the ideal of *truth*. A credence has greater epistemic utility at a world if it's closer to the truth in that world, i.e., if it has greater *accuracy*.

Joyce (1998) offers a non-pragmatic defense of probabilism. His argument appeals to a decision rule that Joyce calls "Dominance."⁴ A credence Cr *strongly* dominates another Cr' iff the accuracy (U) of Cr is greater than the accuracy of Cr' at every world (i.e. iff $U(Cr, w) > U(Cr', w)$ for all $w \in \mathcal{W}$). (Cr *weakly* dominates Cr' iff the $U(Cr, w) \geq U(Cr', w)$ for all $w \in \mathcal{W}$ with inequality in at least one case.) The decision rule Dominance is a prohibition against adopting strongly dominated credences. Joyce then proves that for a range of epistemic utility functions satisfying certain constraints, any credence function that violates probabilism is strongly dominated.

Instead of Dominance, Greaves & Wallace (2006) and Leitgeb & Pettigrew (2010a,b) argue for Bayesian norms by appeal to the decision rule Maximize Expected (Epistemic) Utility. Greaves & Wallace use this norm to defend conditionalization. They argue that, upon receiving new evidence, the update procedure that uniquely maximizes expected accuracy is to conditionalize on the new evidence. Leitgeb & Pettigrew provide expected-accuracy-based arguments for both probabilism and conditionalization.

What all of these arguments have in common is that they aim to derive results simply by plugging into the familiar decision-theoretic apparatus utility functions

by it.

³ Talk of epistemic "options" or "acts" makes it sound as though epistemic utility theory is committed to doxastic voluntarism. But these descriptions are only meant to make the analogy with practical decision theory explicit. The project can be thought of as a system of third-personal evaluations rather than first-personal guides to "action," and so there's no more need for doxastic voluntarism than anywhere else in epistemology.

⁴ I'll call the decision rule "Dominance" with a capital D and the relation "dominance" with a lower

that encode what's epistemically valuable (namely, accuracy) and epistemic acts (adopting credences).

What I'll argue in this paper is that the norms that epistemic utility theorists have appealed to differ in surprising ways from those used in practical decision theory. These norms use different notions of expected utility and dominance from those used in practical decision theory. But so far, the difference hasn't been acknowledged or justified. As I'll show, the differences aren't innocuous. First, they produce very different results; different epistemic actions are sanctioned. Second, without the clear intuitive motivations of the old norms, the new norms have no clear motivation at all.

2 When the world isn't independent of our beliefs

One of the Bayesian norms that epistemic utility theory set out to justify was conditionalization. Greaves & Wallace (2006) and Leitgeb & Pettigrew (2010b) argue that the norm Maximize Expected Accuracy always requires agents to conditionalize on any new evidence:

Maximize Expected Accuracy (MaxExAcc): rational agents have the credences that maximize expected accuracy (from the agent's perspective).

I'm going to argue that on the most natural interpretation of this norm, it conflicts with conditionalization.

The kind of cases where MaxExAcc conflicts with conditionalization are cases where credal acts are not causally or evidentially independent of events in the world.⁵ There are a variety of such cases:

- Your credence in a proposition can causally affect the likelihood that the proposition is true: a belief might be self-verifying,⁶ self-falsifying,⁷ etc.

case d.

⁵ See (Greaves, manuscript), (Caie, forthcoming).

⁶ William James discusses cases like this as examples of practical reasons for belief: "There are then cases where faith creates its own verification. Believe, and you shall be right, for you shall save yourself; doubt, and you shall again be right, for you shall perish" (James, 1896, 96–7).

⁷ The TV show *Arrested Development* provides one such example: an out-of-work actor auditioning for the role of a frightened inmate is more likely to get the role if he believes he won't (because he'll exhibit genuine fear) and less likely to get the role if he believes he will (because he'll fail to exhibit genuine fear).

- Your credence in a proposition can also causally affect the likelihood that other propositions are true.⁸
- Your lower-order credences determine the accuracy of your higher-order credences.
- Similarly for other mental states: your credences about what you will do are at least not evidentially independent of whether they're true.⁹
- Your credences are at least not evidentially independent of others' credences, including their credences about your credences, e.g., in communication.
- Similarly, your credences are not evidentially independent of other facts about the world, e.g. others' actions when you believe you'll command them.
- Your credences in certain propositions can logically determine whether those propositions are true.¹⁰
- Your credences may determine the likelihood of certain kinds of social facts: for example, if I'm The Trendsetter and I believe something is cool, that can make it the case that it is cool (constitutively, rather than causally).

And so on. In short: which world you're in is partly dependent on your epistemic acts. And so we need to take into account this kind of dependence in epistemic utility theory.

First, let's examine the effects of this kind of dependence on the standard machinery of epistemic utility theory. I will be focusing on examples of the first

⁸ See Berker (2013) for examples and discussion.

⁹ Typically you believe you'll do what you intend to do; this belief is, on some views, constitutive of the intention.

¹⁰ For example, in the case of certain self-referential beliefs: the proposition *I believe this proposition is true* is true just in case I believe that proposition; by believing it I make it true. For a discussion of epistemic rationality and self-referential attitudes of this sort, see (Caie, 2012) and (Caie, forthcoming). Caie (forthcoming) focuses on Joyce's dominance argument for probabilism. He argues that Joyce ignores the fact that dominance reasoning is not valid in cases where acts are not independent of the states of the world. Caie endorses the move to some form of causal or evidential epistemic decision theory. In section 3 of this paper, I explain why both causal and evidential versions of epistemic decision theory lead to bad (anti-evidentialist) results precisely in cases where the states are not independent of credal acts. In my view, then, Caie succeeds in showing that dominance arguments (on the traditional decision-theoretic version of dominance) conflict with probabilism. But rather than casting doubt on probabilism, I think this impugns dominance arguments, for reasons I'll explain in section 6.

and second kind.¹¹ In the following example, an agent's credence in a proposition causes that proposition to be more or less likely to be true. I'll return to this example throughout this paper.

Example #1

Suppose your (perfectly reliable) yoga teacher has informed you that the only thing that could inhibit your ability to do a handstand is self-doubt, which can make you unstable or even hamper your ability to kick up into the upside-down position. The more confident you are that you will manage to do a handstand, the more likely it is that you will, and vice versa.

Let's make things more precise: H is the proposition that you'll successfully do a handstand at t_3 . Your yoga teacher has informed you that for all n , $Cr_1(H) = n$ at t_1 will make it the case that $Ch_2(H) = n$ at the next moment, t_2 , and that this will remain the chance of H up to t_3 , when you either do or don't do a successful handstand. We'll call the information she's given you: " $Cr(H) = Ch(H)$," where " Cr " and " Ch " nonrigidly pick out, respectively, whichever credence function you adopt at a world and the chances at that world.

Suppose that in the actual world, your prior Cr_0 is such that before learning $Cr(H) = Ch(H)$, your conditional credence in H given that information is .5:

$$Cr_0(H \mid Cr(H) = Ch(H)) = .5$$

And let's assume for the moment that this is a rational credence to have. After all, the result that conditionalization always uniquely maximizes expected accuracy is supposed to hold for *any* probabilistic priors. So if there's is a divergence between

¹¹ Examples of the first and second kind are central to Berker's (2013) argument against various forms of epistemic consequentialism, such as reliabilism. Suppose an agent's believing that she will recover from a particular illness makes it the case that she will recover from the illness, but she has evidence that 80% of people with this illness do not recover. Is it epistemically required for her to believe that she will recover from the illness, given that doing so is conducive to having true belief? Intuitively not. The focus of Berker's argument is various forms of *objective* epistemic consequentialism. Berker doesn't direct his argument explicitly toward epistemic utility theory, which at first glance seems to be a form of *subjective* epistemic consequentialism. As we'll see, constructing examples of causal dependence between belief and truth for epistemic utility opens up foundational questions about what sort of project epistemic utility theory is really engaged in, and whether it is genuinely a form of subjective epistemic consequentialism.

conditionalization and MaxExAcc, then it doesn't matter whether these priors are rational.

What credence is rational for you to adopt upon learning $Cr(H) = Ch(H)$?¹² Conditionalization, of course, says that your updated credence should be .5. The question is, what do MaxExAcc and Dominance say? We'll turn to MaxExAcc first and return to Dominance in section 6.

3 Conditionalization and expected accuracy

3.1 Two kinds of decision theory

In order to know what MaxExAcc recommends, we need to specify what the rule says. MaxExAcc is a special case of a decision-theoretic rule, Maximize Expected Utility, paired with a particular form of utility: accuracy. There are two competing forms that the rule Maximize Expected Utility most commonly takes in decision theory: Maximize Causal Expected Utility and Maximize Evidential Expected Utility. As we will see, both of these rules are different in interesting ways from the decision rule used in epistemic utility theory, and as a result they make different predictions from the decision rules epistemic utility theorists use. But because epistemic utility theorists have not discussed this difference, it's helpful to see what would happen if we simply took traditional practical decision theory and fed into it an epistemic utility function.

The **causal expected utility** of an act is a measure of the value you can expect to result from taking that act. Here's a natural way of calculating causal expected utility (from Lewis 1981): we partition the space of possibilities into a set of so-called dependency hypotheses. A dependency hypothesis is a maximally specific proposition about how facts about the world causally depend on the agent's present acts. The agent might be uncertain about what effects her acts can have, and so there may be many epistemically possible dependency hypotheses. The causal expected utility is the weighted average of the value of each possible causal outcome of the act, weighted by the probability of the dependency hypothesis where the act causes that outcome. Formally, the causal expected utility is calculated as follows: where each k_i is a possible dependency hypothesis, $CEU(a) = \sum_i Cr(k_i)U(k_i \wedge a)$.

¹²This example has a confounding factor: that it's desirable to be able to do a handstand. This makes the idea that one should have credence 1 in H appealing from a pragmatic perspective. Indeed, James (1896) discusses examples of this sort of belief efficacy as a way of arguing that there can be practical reasons for belief. There are (slightly less realistic) examples without this confounding factor: for example, suppose your beliefs determine whether some freckle will appear on your right arm rather

The **evidential expected utility** of an act differs from its causal expected utility in roughly this way: it doesn't distinguish between cases where your acts are likely to cause a certain outcome and cases where your acts merely correlate with outcomes (perhaps because, e.g., they have a common cause, or perhaps just by chance). The evidential expected utility of an outcome is calculated as follows: where each s_i is a possible state of the world (such that the set of states forms a partition), $EEU(a) = \sum_i Cr(s_i | a)U(s_i \wedge a)$.

There are other forms of decision theory, but these two are the most familiar forms and I'll focus on them in the remaining discussion.

3.2 Causal expected accuracy

Returning to our example: for simplicity's sake, let's suppose once you've learned that your credence in H determines the chance of H such that $Cr(H) = Ch(H)$, there are only three options for which credence function to adopt¹³:

$$\begin{aligned} Cr_a(H) &= 1 \\ Cr_b(H) &= .5 \\ Cr_c(H) &= 0 \end{aligned}$$

We'll assume that the only credence functions you can adopt are probability functions that give credence 1 to $Cr(H) = Ch(H)$. So once we stipulate that the only propositions these credence functions are defined over are H , propositions concerning which credence function you have, propositions concerning which chance function obtains, and the Boolean closure of all of them, then this is all we need to fully specify each credence function.

In our example, then, we only have to consider two dependency hypotheses,

than your left, or determine whether it'll rain on some uninhabited planet.

¹³ I'll leave it as an exercise to the skeptical reader to fill in the remaining options $Cr(H) \in (0, .5) \cup (.5, 1)$; these other options won't make any difference to my point.

which differ only in whether H would be true if you were to adopt Cr_b .¹⁴

$$\begin{aligned}
k_1 : & \text{ you adopt } Cr_a \Box \rightarrow (Ch(H) = 1 \text{ and } H) \\
& \text{ you adopt } Cr_b \Box \rightarrow (Ch(H) = .5 \text{ and } H) \\
& \text{ you adopt } Cr_c \Box \rightarrow (Ch(H) = 0 \text{ and } \overline{H}) \\
k_2 : & \text{ you adopt } Cr_a \Box \rightarrow (Ch(H) = 1 \text{ and } H) \\
& \text{ you adopt } Cr_b \Box \rightarrow (Ch(H) = .5 \text{ and } \overline{H}) \\
& \text{ you adopt } Cr_c \Box \rightarrow (Ch(H) = 0 \text{ and } \overline{H})
\end{aligned}$$

The reason we only have these two is that we're including the information that $Cr(H) = Ch(H)$. Other dependency hypotheses where, e.g., $Cr_a \Box \rightarrow \overline{H}$ are all automatically given probability zero, and so we don't have to consider them.

So there are two states, k_1 and k_2 , and three options. This gives us a partition over epistemic possibilities; we can treat the cells of the partition as worlds. We then calculate the accuracy of each epistemic act at all worlds where that act is taken.

We're interested in the local inaccuracy of your credence in a particular proposition.¹⁵ A **local inaccuracy** measure characterizes the inaccuracy of an agent's credence in a particular proposition. A **global inaccuracy** measure characterizes the inaccuracy of a total credence function. As an inaccuracy measure, we'll just use the negative Brier score.¹⁶ Let $v_w(\cdot)$ be the omniscient probability function at a world w (mapping all truths to 1 and falsehoods to 0). Then the negative Brier score of $Cr(X)$ at w is:

$$U(Cr(X), w) = -(v_w(X) - Cr(X))^2$$

Now we can assign values to the worlds in our decision problem:

¹⁴ Notation: " $\Box \rightarrow$ " is the counterfactual connective: *if... would...* I've represented the dependency hypotheses as including determinate facts about H . Lewis (1981), by contrast, does not distinguish dependency hypotheses that differ only with respect to chances, and so would treat this as a case where there's only one dependency hypothesis with positive probability. The end result is the same in this case; I prefer the non-Lewisian variation because it allows full specification of the values of outcomes even if they're chancy. We can't say how accurate a credence in H is at a state unless we specify whether H is true at that state.

¹⁵ We can get away with this because in our toy example, the global inaccuracy of a credence varies directly with the local inaccuracy of the credence in H .

¹⁶ The Brier score is widely accepted as a good inaccuracy measure; see (Joyce, 2009) and (Leitgeb & Pettigrew, 2010a) for philosophical defenses of it. It's negative because I've been talking about maximizing accuracy, rather than minimizing inaccuracy. Ideally accurate credences have a Brier score of 0.

	k_1	k_2
Cr_a	0	0
Cr_b	-.25	-.25
Cr_c	0	0

There's no need to calculate the causal expected accuracy because Cr_b is strongly dominated.¹⁷ So we know that Cr_b is impermissible, according to causal MaxEx-Acc.

Causal MaxExAcc therefore conflicts with conditionalization. Conditionalization requires you to adopt Cr_b , whereas causal MaxExAcc requires you to adopt either Cr_a or Cr_c .

3.3 Evidential expected accuracy

Now we'll figure out what evidential MaxExAcc recommends. First, note that only four possibilities (w_1, w_2, w_3, w_4) have positive probability.

	Ch_a, H	Ch_b, H	Ch_c, H	Ch_a, \bar{H}	Ch_b, \bar{H}	Ch_c, \bar{H}
Cr_a	w_1	/	\	\	/	/
Cr_b	/	w_2	\	\	w_3	/
Cr_c	/	/	\	\	/	w_4

\ = ruled out by chance info; / = ruled out by $Ch(H) = Cr(H)$

The evidential expected accuracy calculation is as follows. Notation: "Cr", in sans serif, denotes the proposition that you adopt the credence function Cr ; "w", in

¹⁷ By "dominated," here I mean the term in the sense used in practical decision theory: that there's some alternative act such that for every state (that is, dependency hypothesis), the outcome of Cr_b is worse than every outcome of performing that act. As we'll see in section 6, epistemic utility theorists have used a different definition of "dominance."

sans serif, denotes the maximally specific proposition that is only true at w .¹⁸

$$\begin{aligned}
EEA(Cr_a(H)) &= Cr_0(w_1 \mid Cr_a)(-(v_{w_1}(H) - Cr_a(H))^2) \\
&= 0 \\
EEA(Cr_b(H)) &= Cr_0(w_2 \mid Cr_b)(-(v_{w_2}(H) - Cr_b(H))^2) \\
&\quad + Cr_0(w_3 \mid Cr_b)(-(v_{w_3}(H) - Cr_b(H))^2) \\
&= -.25 \\
EEA(Cr_c(H)) &= 0
\end{aligned}$$

Whatever your prior credence function (Cr_0) is, as long as it's probabilistically coherent and updated on $Cr(H) = Ch(H)$, it will give Cr_b lower evidential expected accuracy than Cr_a or Cr_c . And the same holds for any alternative to Cr_a or Cr_c that assigns non-extremal values to H and \bar{H} . If your credence in H is extremal, the distance between your credence and the truth will be 0. But for any non-extremal credence, there will be some distance between your credence and the truth.

So no matter what your prior is, evidential MaxExAcc instructs you to adopt extremal credences. And this means that, like causal MaxExAcc, it conflicts with conditionalization in all cases where $Cr_0(H \mid Cr(H) = Ch(H))$ is not extremal. Indeed, it's no surprise that causal and evidential MaxExAcc agree in this case: the dependency relation between H and your credences is causal as well as evidential.

3.4 Discussion

Greaves & Wallace (2006) and Leitgeb & Pettigrew (2010b) argued that updating by conditionalization uniquely maximizes expected accuracy, from the perspective of any probabilistic priors. I've shown that if MaxExAcc is cashed out evidentially or causally, this is not true.

Now, one might concede this point and be willing to give up on the general result—but still want to embrace both conditionalization and causal or evidential expected accuracy. And there is a fall-back position that allows one to do so in the example I gave: by ruling nonextremal credences in H irrational. We can't get the right results for *all* probabilistic credences, one might say; but we can at least get it for the rational ones.

¹⁸ Note: $U(Cr, w)$ doesn't use the sans serif notation because U 's arguments are a credence function and a world, not a conjunction of propositions. The reason for this is that any credence function can be paired with any world as input to U , even when $Cr \cap w = \emptyset$ (i.e. whenever w happens to be a world where you adopt a different credence function, or don't exist at all). The fact that the utility functions range over such worlds is, in fact, central to the puzzle this paper discusses.

Here are three reasons for skepticism about this strategy. First, it's intuitively implausible that an ideally rational agent could never have good evidence for thinking: "I'm just the kind of person who'll be uncertain about H in circumstances like this; and so there's no telling whether H will be true." Second, learning $Cr(H) = Ch(H)$ doesn't *necessarily* give any information in favor of or against H . Indeed, it's perfectly symmetrical with respect to whether H is true; why would it be irrational to have a credence like .5, which is also symmetrical about whether H is true? Third, what's rational to believe depends on what the evidence supports. But it would be strange to think that the evidence supported both our hypothesis H to degree 1 and to degree 0—but nothing in between!¹⁹

The better conclusion to draw from these sorts of examples is that the goal of having accurate credences is different from, and can even conflict with, the goal of having the credences that are supported by the evidence. I'll discuss this conclusion more in section 7.²⁰ But this isn't the primary focus of this paper; rather, this paper focuses on what examples like these reveal about some of the central concepts at work in epistemic utility theory: expected accuracy and accuracy dominance. I want to explain in more detail how they differ from the notions of expected utility and dominance in decision theory, and to raise questions about whether the rules derived from them have genuine normative force.

3.5 Expected accuracy in epistemic utility theory

Now, you might think I have provided a counterexample to the proofs in Greaves & Wallace (2006) and Leitgeb & Pettigrew (2010b) that MaxExAcc entails conditionalization. But I haven't. What I've shown is that if we conceive of expected

¹⁹ A natural thought: What if instead of measuring credences' utility by distance from *truth*, we measure them by distance from *objective chance*? Then whatever your credence in H is, it'll have zero distance from the chance; all options are equally good. So we end up with a maximally permissive recommendation for H . And maybe that's an attractive outcome in this case.

Even if you think this is the right result for our first example, there are problems for this account. First, it still doesn't vindicate conditionalization, since it permits you to update to credences that aren't permitted by conditionalization. Second, the solution doesn't generalize. Consider another example: suppose you learn that whatever credence you adopt in H will make it the case that the objective chance of H is .5 lower or higher—*unless* you adopt credence .79 in H , in which case $Ch(H) = .79$. In these circumstances, adopting $Cr(H) = .79$ will maximize expected proximity to objective chance. But this is no reason to think that the evidence supports H to degree .79.

In this paper I'll be mainly focusing on accuracy measures. But I should acknowledge that other types of epistemic utility function could be used to generate a variety of interesting results.

²⁰ See (Berker, 2013) and (Berker, forthcoming) for a more thorough discussion of the difference between evidential support and truth-conducivity.

accuracy in the common decision-theoretic way—as something like causal or evidential expected accuracy—then we can’t get this result.

Evidential expected accuracy: $EEA_{Cr}(Cr') = \sum_{w \in W} Cr(w \mid Cr')U(Cr', w)$

Causal expected accuracy: $CEA_{Cr}(Cr') = \sum_{k \in K} Cr(k)U(Cr', k)$

Philosophers working on epistemic utility theory, including Greaves & Wallace and Leitgeb & Pettigrew, use a notion of expected accuracy different from either causal or evidential expected accuracy. We’ll call their inaccuracy measure “observational expected accuracy” or “OEA”: it measures expected accuracy for pure observers, whose epistemic acts have no dependency relations with facts about the world. Where E is the set of epistemically possible worlds,

Observational expected accuracy: $OEA_{Cr}(Cr') = \sum_{w \in E} Cr(w)U(Cr', w)$

Observational MaxExAcc does recommend the same credal acts as conditionalization.

How? OEA doesn’t take into account, in its calculation of the expected accuracy of an epistemic act, the fact that the epistemic act is taken. That is, it doesn’t take into account any dependency relations (causal or evidential) between adopting particular credences and what the world is like. By contrast, evidential expected accuracy weights the utility of an epistemic act at different worlds by the probability of that world *conditional on your performing that epistemic act*. Causal expected accuracy uses partitions of worlds that are based on the effects your acts might have. In both cases, the weighting on the utilities for the expectation depends on your taking the act in the worlds considered.

The difference between OEA and either of the traditional notions of expected utility has gone largely unnoticed. Those who have noticed this difference (Greaves, manuscript), or a corresponding difference between dominance in decision theory and “dominance” in epistemic utility theory (Caie, forthcoming) have argued in favor of revising epistemic utility theory from the pure observational picture to the decision theoretic (act-sensitive) picture.²¹ But the use of OEA has wide-ranging

²¹ Greaves argues that evidential and causal epistemic decision theories are subject to counterexamples based on cases discussed in practical decision theory (Newcomb-type cases for EEA; instability cases for CEA), and expresses doubt about the prospects for epistemic utility theory, at least insofar as its intended to justify intuitions about epistemic rationality. She does not discuss the conflict between conditionalization and both CEA and EEA, nor probabilism’s conflicts with each, and does not discuss how this might provide a motivation to stick with what I’ve called OEA. Greaves also does not discuss ways in which dominance arguments in expected utility theory rely on very similar simplifying assumptions.

effects on epistemic utility theory.²²

4 Probabilism and expected accuracy

Now I'll briefly show that, with a slightly more complicated example, the argument in the previous section extends to probabilism. Both causal and evidential expected

²² And now, a brief excursus on how, if we switch to CEA or EEA, not only can we not get the same conclusions from epistemic utility theory, but we can't even rely on the same premises. A standard starting point in epistemic utility theory is to characterize epistemic utility functions. Many, including Joyce (1998, 2009) and Leitgeb & Pettigrew (2010a), aim for as much generality as possible and so argue for weak constraints on epistemic utility functions. A constraint on epistemic utility functions that Gibbard (2007) introduced and Joyce and Leitgeb & Pettigrew defend is that they make rational credences *immodest*, in the following sense:

Immodesty: Any rational credence Cr will assign itself higher expected epistemic utility than any other credence function.

As Joyce notes, a variety of different epistemic utility functions satisfy this desideratum. (These epistemic utility functions are called "strictly proper" scoring rules.) This is so, though, only with observational expected accuracy. What happens if we use evidential or causal expected accuracy?

The example we began with is an obvious case where the Brier score is compatible with modest probabilistically coherent credences when expected accuracy is interpreted causally. (I'll leave it an exercise to the interested reader to see why; simply consider the perspective of Cr_b .)

The argument involving EEA is more complicated. We can call a credence function "transparent" iff it conforms to this:

Transparency: $Cr(H) = n$ only if $[Cr(Cr(H) = n) = 1$ and for all $m \neq n$, $Cr(Cr(H) = m) = 0]$.

With either causal or evidential expected accuracy, only credence functions that are transparent can maximize causal or evidential global expected accuracy. (I will explain why in section 5.3.) So, we'll restrict ourselves to transparent credence functions. But then the EEA of all credence functions other than one's own will be undefined. After all, $EEA_{Cr}(Cr')$ is calculated using $Cr(w | Cr')$. But since Cr must be transparent, it assigns Cr' probability zero, and so the conditional probability will be undefined. And the EEA of your own credence function won't be higher than any other credence function: all alternatives will have undefined EEA.

So the Brier score is not a proper scoring rule when expected accuracy is interpreted in the traditional decision-theoretic way! And in fact, any plausible accuracy measure will also be improper: extremal credences in H will perfectly match the truth, and so their local inaccuracy will be 0 by any measure of distance from the truth. By contrast, nonextremal credences won't *perfectly* match the truth, and so their inaccuracy will be positive by any plausible inaccuracy measure.

These sorts of examples show that, combined with ordinary forms of expected utility, immodesty isn't a good epistemic principle. While it's plausible that rational credences should assign themselves some privileged status—for example, perhaps they should consider themselves the only rational response to the evidence at hand, or more likely to be rational than any particular alternative—they can sometimes rationally assign themselves lower expected accuracy than some alternative.

accuracy maximization will, under some circumstances, require probabilistically incoherent credences.

Example #2

Your perfectly reliable yoga teacher has informed you that at a future time t_1 , your credence in H will causally determine the chance of H (so that at the next moment, t_2 , the chance is set until H does or does not turn out to be true at t_3). This time, your credence will always be .5 greater or less than the chance:

$$\begin{aligned} \text{if } Cr_1(H) > .5, \text{ then } Ch_2(H) &= Cr_1(H) - .5, \text{ and} \\ \text{if } Cr_1(H) \leq .5, \text{ then } Ch_2(H) &= Cr_1(H) + .5. \end{aligned}$$

So you know at t_0 that whatever credence you adopt at t_1 , it'll be far from the objective expected accuracy (i.e. the chance-weighted expectation of accuracy). The oracle then tells you that there is one exception: if your credences are such that $Cr_1(H) = 1$ and $Cr_1(\neg H) = .1$, then H is true. We'll call this credence function Cr_a .

Now, whether your credences are probabilistic or not, if you don't adopt Cr_a , then the upper bound on your accuracy (measured here by the sum of the negative Brier scores of $Cr(H)$ and $Cr(\neg H)$) is $-.25$. But if you adopt Cr_a , then your expected accuracy is $-.01$.

Hence, the option of adopting Cr_a has maximum expected accuracy. But Cr_a is probabilistically incoherent.

One might conclude from examples of this sort that probabilistically incoherent credences are sometimes rationally required. For example, Caie (forthcoming) argues for this conclusion: since there are circumstances where we can get our credences closer to the truth by violating probabilism, we sometimes rationally ought to violate probabilism.

I claimed in the introduction that epistemic utility theory faces a dilemma: we must adopt a decision theory that reflects how actions can affect how the world is (thereby giving up conditionalization and probabilism), or else explain what OEA is and why we should maximize it. To say that maximizing OEA ensures that our beliefs are closer to the truth can't be right. Examples #1 and #2 are counterexamples: in both cases the rational agent can be certain that credences that violate observational MaxExAcc are closer to the truth.

I interpret Greaves and Caie as accepting the first horn of the dilemma. But I think examples like those I've discussed push in favor of assessing the prospects

for the second horn. If rationality requires having the credences that the evidence supports, then examples #1 and #2 show how rationality doesn't require minimizing our credences' expected distance from the truth.

5 What is observational expected accuracy?

5.1 OEA's idealizing assumption

Maximizing OEA, we are told, is rational because it allows us to approach the ideal of having credences that are as accurate as possible (given our limited information). But this is only true on the assumption that an agent's belief state is always causally or evidentially independent of the propositions the agent has beliefs about. In section 2, I showed that far from being oddball cases, counterexamples to this assumption are common. In these counterexamples, getting our credences as close as possible to the truth can require choosing options that fail to maximize OEA. Why do epistemic utility theorists rest content with this strong idealization?

One of the background presuppositions might be that the "worlds" don't distinguish epistemic acts. The world is entirely outside of us; we are causally and evidentially inert observers. Hence there are no dependency relations between beliefs and the rest of the world; hence there is no reason for higher-order credences. OEA runs into trouble as soon as we allow for credal actions to be parts of the world, which we can have credences in.

I think this is an instance of a wider phenomenon. Sometimes philosophers working in epistemic utility theory ignore or idealize away the fact that some of the propositions that probability functions range over are propositions about probabilities. In other words, probabilities (subjective and objective) iterate.²³

²³ For example: Joyce (2009) proposes that for any measure to be an inaccuracy measure, it must necessarily rule that every probability function is not accuracy dominated. The reason is that he thinks every probability function could in principle be rational by the lights of some body of evidence, in particular evidence about chances:

[F]or any assignment of probabilities $\langle p_n \rangle$ to $\langle X_n \rangle$ it seems that a believer could, in principle, have evidence that justifies her in thinking that each X_n has p_n as its objective chance. (279)

The thought: chances form a probability function. For any probability function, you could have evidence that entailed that the chances of each proposition were just those of that probability function. And so, with the principal principle, the rational credences would also form a probability function.

Joyce's argument is only plausible if the believer's credence function doesn't range over chance facts, and so we can ignore the fact that chances iterate. After all, there are probabilistically coherent credence functions that have the property of assigning $Cr(H) = n$ and $Cr(Ch(H) = n) = 0$. And these

The big question is: once we move beyond this idealization, does epistemic utility theory give us the right results? Any epistemic utility theory that says to maximize expected accuracy in a way that takes into account the effect of epistemic acts on the world (like EEA or CEA) will no longer uphold conditionalization or probabilism. On the other hand, observational MaxExAcc does uphold conditionalization and probabilism—but, as I’ll argue, OEA isn’t genuinely a measure of the expected accuracy of credal acts.

5.2 The standard motivations for MaxExAcc don’t motivate maximizing OEA

Because the differences between OEA and both CEA and EEA have been widely ignored, the intuitive justifications that have been given for maximizing OEA have pointed in the wrong direction. They do nothing to justify maximizing OEA and instead motivate norms like maximizing CEA or EEA: norms that require adopting the credences that best approximate the truth *in the worlds where you have those credences*. In other words, CEA and EEA maximization, unlike OEA maximization, are norms that require taking credal actions that can be predicted to lead to good epistemic outcomes.

Let me list some examples of considerations intended to support maximizing OEA that actually support maximizing CEA or EEA (or similar norms), and in some cases are not even consistent with maximizing OEA. (As we’ll see in section 6, there’s an analogous distinction between dominance in decision theory and “dominance” in, e.g., Joyce’s accuracy dominance arguments. The quotations below are general enough to justify both the traditional decision theoretic interpretations of expected accuracy and of dominance.)

Here is Joyce’s defense of a norm of gradational accuracy:

My position is that a rational partial believer must aim...to hold partial beliefs that are gradationally accurate by adjusting the strengths of her opinions in a way that best maximizes her degree of confidence in truths while minimizing her degree of confidence in falsehoods. (Joyce, 1998, 578)

The fact that one set of credences incurs a lower [inaccuracy score]

can even be perfectly justified. For example, an agent might be sure a coin is biased either heads or tails, and so have the credence $Cr(Ch(\text{heads}) = .5) = 0$. The same agent might not know which bias the coin has, and so have $Cr(\text{heads}) = .5$. But then it can’t be the case that the believer’s evidence justifies her in thinking that the chance of every proposition X_n is equal to her credence in X_n . That would entail that the chance of heads is .5 and the chance of $Ch(\text{heads}) = .5$ is 0. That’s clearly false.

than another at a given world should be taken to mean that, from a purely epistemic point of view, it would be better in that world to hold the first set of credences than to hold the second. (Joyce, 2009, 266)

Greaves & Wallace's justification for maximizing accuracy is straightforwardly incompatible with OEA:

[C]onditionalization will be epistemically rational if and only if it can reasonably be expected to lead to epistemically good outcomes. (Greaves & Wallace, 2006, 608)

[I]t is (presumably) epistemically better to have higher credences in truths and lower credences in falsehoods. According to the cognitive decision-theoretic approach, epistemic rationality consists in taking steps that can reasonably be expected to bring about epistemically good outcomes. (Greaves & Wallace, 2006, 610)

Similarly for Gibbard:

When a person forms her credences with epistemic rationality, our hypothesis will now run, it is as if she were voluntarily choosing her credences with the pure aim of truth—that is to say, to maximize the expected accuracy of her credence. (Gibbard, 2007, 149)

Leitgeb & Pettigrew are most naturally interpreted as endorsing a norm of having true beliefs or having credences that approximate the truth in the worlds where you have those beliefs:

An epistemic agent ought to approximate the truth. In other words, she ought to minimize her inaccuracy. (Leitgeb & Pettigrew, 2010a, 202)

It is often said that the epistemic norms governing full beliefs are justified by the more fundamental epistemic norm *Try to believe truths*. . . We will appeal to the more fundamental norm *Approximate the truth*, which is plausibly the analogue of the fundamental norm for full beliefs stated above. (Leitgeb & Pettigrew, 2010b, 236–7)

What these quotations have in common is that they can only naturally be interpreted as saying that rational beliefs should be formed with the aim of accuracy. But OEA does not vindicate this intuition. As we saw in the handstand example, OEA will sometimes recommend performing credal acts that a rational agent can be *certain* will leave her farther from the truth than some alternative. The same, we'll see in section 6, is true of observational MaxExAcc's analogue, observational Dominance. But first, let's see why OEA has this sort of result.

5.3 The truth independent of our credences?

Now, here is a natural defense one might give of OEA. CEA and EEA generate conflicts with conditionalization precisely *because* they reflect the causal and evidential dependency relations (respectively) between our credences on the world. What we should aspire to believe, this argument goes, is what's true *independent of our beliefs*. And OEA allows us to measure the expected proximity of our belief states to this independent truth.

The problem with this response is that there are a variety of ways in which, intuitively, our beliefs do affect what it's rational for us to believe. For example, if I know $Cr(H) = Ch(H)$ and I know $Cr(H) = .4$, then I should infer $Ch(H) = .4$. But that fact about chance isn't independent of my beliefs.

In other cases, it's not entirely clear to what degree rationality requires us to take into account our own beliefs: in particular, with respect to higher-order beliefs.

To make the point simply, let's begin with a toy example, where the relevant acts aren't pure epistemic acts. Suppose you're currently uncertain about B , the proposition that you'll go to the beach: $Cr_0(B) = .5$. You have four options: to go to the beach or not go to the beach, and to be certain about whether you'll go to the beach or retain uncertainty.

	B	\bar{B}
$a_1: B \wedge Cr(B) = 1$	w_1	/
$a_2: \bar{B} \wedge Cr(B) = 0$	/	w_2
$a_3: B \wedge Cr(B) = .5$	w_3	/
$a_4: \bar{B} \wedge Cr(B) = .5$	/	w_4

It's clear that a_1 and a_2 —going to the beach and being certain that you are, or not going to the beach and being certain that you aren't—involve adopting more accurate credences than a_3 and a_4 . Whichever of the first two options you take, it will involve adopting a credence in B that is maximally accurate. You can know in advance that either option will lead to greater accuracy. So they should also have greatest expected accuracy.

But OEA predicts that a_3 and a_4 will have higher expected accuracy.²⁴ From a decision-theoretic perspective, this is a bizarre result. It's not particularly strange to think that it might sometimes be rational to be uncertain about whether you're at the beach. But it is strange to think that *on the supposition that you go to the beach*, a .5 credence in B has a higher expected accuracy than credence 1.

Of course, an option like $B \wedge Cr(B) = 1$ isn't a pure epistemic act. But this case can naturally be transformed into a case involving only pure epistemic acts. Using the same example, reinterpret ' B ' as designating the proposition that $Cr(H) = 1$. In other words, make the case one where the options consist of pairs of a first- and second-order credence (where the latter is about the former).

Again, it's clear that the local accuracy of the higher-order credence in a_1 and a_2 is greater than the local accuracy of the higher order credence in a_3 and a_4 . Whether the expected accuracy of the first-order credence is maximal or not, one of the first two options will certainly have greater global accuracy than all of the rest of the options. After all, they guarantee you a free true belief.

Maximizing OEA requires taking one of the latter two options. But if you perform one of the latter two options, you can be certain of at least one particular alternative option that it will land you greater accuracy. If you do what OEA recommends, you are declining a free true belief! And this will be true of any credal act that is not transparent:

Transparency: $Cr(H) = n$ only if [$Cr(Cr(H) = n) = 1$ and for all $m \neq n$, $Cr(Cr(H) = m) = 0$].

So OEA doesn't seem like it's really measuring the expected accuracy of performing credal acts at all.

Furthermore: in examples of this form, maximizing OEA can require having Moore credences. By "Moore credences" I mean credences such that it would be natural to say of the agent that they believe H and believe that they don't believe H . Without taking a stand on the relation between belief and credence, we can offer $Cr(H) = 1$ and $Cr(I \text{ don't believe } H) = 1$ as a natural candidate for such credences. If the person asserted both beliefs as a conjunction, they would assert a Moore sentence: " H is true and I don't believe H ."

It's clear of any credence function that includes Moore credences that it is less accurate than a transparent alternative with the same first order credences. While the person who does indeed adopt Moore credences cannot realize it, this fact is

²⁴ The partition of states that OEA uses is $\{w_1, w_2, w_3, w_4\}$. Let's assume your prior is maximally

clear from a third-person perspective: an assessor who has precisely the same information as the agent about both H and whether the agent believes H will see that a person with Moore credences will be *necessarily* farther from the truth than a person with transparent credences. Only transparent credal acts *should* satisfy MaxExAcc. (This isn't to say that in order to be *rational* you have to have transparent credences; perhaps we can have rational doubts about our own beliefs.)

The fact that maximizing OEA requires sometimes adopting Moore credences is a bizarre result, from a decision-theoretic perspective. Part of the justification that Joyce (2009) gives for why our credences should maximize expected accuracy by their own lights involves an appeal to the irrationality of believing Moore sentences:

If, relative to a person's own credences, some alternative system of beliefs has a lower expected epistemic disutility, then, by her own estimation, that system is preferable from the epistemic perspective. . . . This is a probabilistic version of Moore's paradox. Just as a rational person cannot fully believe "X but I don't believe X," so a person cannot rationally hold a set of credences that require her to estimate that some other set has higher epistemic utility. (277)

Joyce does not acknowledge that if you have probabilistically coherent Moore credences, then you will maximize OEA from your own perspective with any strictly proper scoring rule. After all, there's no conflict between Moore credences and probabilism. To avoid this problem, it isn't sufficient to limit the kinds of accuracy measures we use as epistemic utility functions. The problem isn't in the choice of epistemic utility functions: it's in OEA.

5.4 Expected accuracy of credal acts

In the rest of this section I'm going to try to motivate the following hypothesis: there are necessary conditions on being the expected accuracy of a person's credences, or a credal act, that OEA doesn't satisfy. I will propose two candidate necessary conditions for being the expected accuracy of a person's credences, or of a credal act, and show that OEA doesn't satisfy them.

Superdominance

Say that an option a_1 strongly superdominates an option a_2 iff all possible outcomes of a_1 are better than all possible outcomes of a_2 . In other words, a_1 strongly super-

uncertain about which you'll perform. Then $OEA(a_1) = OEA(a_2) = \sum_w -.25(v_w(B) - Cr_{a_1}(B))^2 =$

dominates a_2 iff for all s, s' in the set of states of the world, $U(s, a_1) > U(s', a_2)$. (a_1 weakly superdominates a_2 iff for all s, s' in the set of states of the world, $U(s, a_1) \geq U(s', a_2)$ with strict inequality in at least one case.)

It seems to me pretty plausible that strongly superdominated options cannot maximize expected utility on any reasonable conception of expected utility. I want to suggest that this is a necessary condition for counting as a measure of expected utility: if an option maximizes something but that option is strongly superdominated, then whatever that option maximizes is not expected utility.

Note: we wouldn't want to say the same thing about ordinary strong dominance, because evidential decision theory allows that strongly dominated options can sometimes maximize expected utility. Superdominance, on the other hand, is something that causal and evidential expected utility can agree on.

Now, no option that maximizes CEA or EEA is strongly superdominated. If every outcome of a_1 has greater utility than every outcome of a_2 , then it doesn't matter how you do a weighted average of the utilities: a_1 will have greater expected utility than a_2 . But there are cases where strongly superdominated options maximize OEA.

For example: in the handstand case, adopting credence 1 in H will give you a Brier score of 0 (i.e. a perfect score) at all epistemically possible states of the world. Adopting credence .5 in H will give you a Brier score strictly less than 0 at all states. But from the perspective of your priors, the OEA of credence .5 in H is greater than the OEA of having credence 1 in H . So for these two options, maximizing OEA requires adopting a strongly superdominated option.

Known accuracy

This example points to a second candidate necessary condition on being a measure of expected accuracy of credal acts that OEA doesn't satisfy. If the value of a random variable is known, then from the perspective of the knower, its expected value should be equal to its actual value.

Consider once more the handstand example. You know in advance that if you adopt $Cr(H) = 1$, the (local) accuracy your credences will have is 0. But the (local) OEA of credence 1 in H , from your point of view, will not be 0. The only assumption we need to make about the accuracy measure is that a credence in a proposition matches the proposition's truth value iff its accuracy is 0 (i.e. perfect; its distance from the truth is 0). Then, in examples like this, the known accuracy of a credal act will not match its OEA.

$$-.5 \text{ and } OEA(a_3) = OEA(a_4) = \sum_w -.25(v_w(B) - Cr_{a_3}(B))^2 = -.25.$$

So, if this is a necessary condition on expected accuracy of a person's credences, then OEA does not tell us the expected accuracy of a person's credences.

What is OEA?

On my view, this shows that OEA isn't the expected accuracy of a person's credences, or credal acts. It stands in need of a philosophical interpretation. Without one, it can't do the work it's been thought to do in epistemology. While it's at least intuitive that maximizing expected accuracy might be thought of as a good thing to do epistemically, it's not at all clear why maximizing *whatever OEA is* is good. So it's hard to see how maximizing OEA, whatever it is, can justify any other epistemic norm.

What we can say of it is this: if epistemic utility theory is meant to be bringing decision theory to epistemology, then we should be concerned with the epistemic value people's credences could have if they take particular credal acts. OEA isn't concerned with people's credences or credal acts. It's concerned with credence functions conceived as abstract objects: functions that exist at every possible world, whether or not they characterize anyone's doxastic state at any given world.

This characterization of OEA raises a number of questions: why should we be concerned with the closeness of an abstract object to the truth, but not with the closeness of our own beliefs to the truth? What does it mean to assign them epistemic value? Why should are worlds where no one takes a certain credal act relevant to the assessment of that credal act?

As we'll see in the next section, questions of the same form carry over to Joyce's accuracy-"dominance" argument for probabilism.

6 Problems for accuracy dominance arguments

6.1 Joyce's "dominance" argument

In the discussion of superdominance, you might have wondered: didn't Joyce (2009) prove that, for a certain set of accuracy measures including the Brier score, no probability functions are accuracy dominated? Strong superdominance entails dominance, and so if probabilistically coherent credence functions are never dominated, then they are never superdominated.

The relation that Joyce calls "dominance" in (Joyce 1998, 2009) is different from the familiar notion of dominance. Let's call Joyce's version "observational dominance" or "O-dominance." Here's how he defined the relation of O-dominance:

Cr [O-]dominates Cr' relative to utility function U iff $U(Cr, w) > U(Cr', w)$ for all $w \in W$.

By contrast, here is the standard definition of dominance in practical decision theory:

Cr strongly dominates Cr' relative to U iff $U(Cr, s) > U(Cr', s)$ for all $s \in S$.

The disanalogy between the dominance and O-dominance is clear: dominance only considers the value of an act at outcomes where that act is performed. But O-dominance considers the value of an act at outcomes where that act isn't performed.

But what is the value of an act at an outcome where the act isn't taken? What does that even mean?

There's no difficulty in carrying out the mathematical procedure of applying a scoring rule to a credence function at a world where no one adopts that credence function. But what exactly are we measuring by doing so? It doesn't make sense to talk of an act's value at a world where it isn't performed. Certainly, it makes sense to talk of the closeness of a credence function, understood as a mathematical object, to the truth values at a world where you don't take that credence function. But we cannot interpret that as the *epistemic utility of adopting that credence function* at that world, or the *accuracy your credences would have if you adopted that credence function* at that world.²⁵

Now, because O-dominance considers the accuracy of your epistemic acts at worlds where you don't take them, the set of worlds under consideration for O-dominance is a superset of the set of worlds under consideration for dominance. So the claim that Cr O-dominates Cr' is a stronger claim than the claim that Cr dominates Cr' . In some cases, Cr dominates but doesn't O-dominate Cr' . And the examples I gave in the previous section (where probabilistically coherent credence functions were superdominated) were cases where those credence functions were dominated, but not O-dominated.

Joyce's (1998, 2009) "dominance" argument for probabilism is an O-dominance argument for probabilism. A premise of his argument is that it is irrational to adopt credences that are accuracy-O-dominated. But the rationale for that equally motivates the claim that it's irrational to adopt credences that are accuracy dominated. Here is Joyce's (2009) justification for the claim that it's irrational to adopt O-dominated credences:

²⁵ Note that causal decision theory (but not evidential decision theory) allows you to calculate the ex-

[W]hen we endorse a rule as the correct measure of epistemic disutility we commit ourselves to thinking that there is something defective, from a purely epistemic perspective, about credences that score poorly according to that rule. Moreover, if these poor scores arise as a matter of necessity, then the defect is one of epistemic irrationality. (267)

Joyce claims that O-Dominance is not a substantive thesis about epistemic rationality, but rather a constraint on what can be considered an epistemic utility function. But the claim that “if poor scores arise as a matter of necessity, then the defect is one of epistemic rationality” justifies Dominance at least as well as O-Dominance. After all, adopting dominated credences means adopting credences that are farther from the truth than some alternative, no matter what the rest of the world is like. And so requiring rational credences not to be O-dominated is at the very least substantive. What justifies the rule O-Dominance over Dominance?

If, however, we use the decision-theoretic conception of dominance, then there will be some probabilistically coherent credence functions that are dominated. Furthermore: it’s not obvious that all credences that are accuracy-dominated are irrational. If it’s ever rationally permissible to be uncertain about what your credences are, then in those circumstances it’s rationally permissible to have accuracy-dominated credences. And, of course, there are other ways in which credal acts affect outcomes. Suppose we look at dominance within the set of epistemically possible worlds. In the handstand example, when you learn $Cr(H) = Ch(H)$, some transparent, probabilistically coherent credence functions with extremal credence in H will dominate transparent, probabilistically coherent credence functions with non-extremal credence in H . But in those circumstances, intuitively, nonextremal credence in H is rationally permissible.

So: we need some justification for why choosing options that are O-dominated would be irrational, given that choosing options that are dominated is not. But this justification cannot be presented in terms of the aim of having accurate credences. If your aim is accuracy, why does it matter whether your credences are only closer to the truth than some dominating alternative credences at worlds *where you don’t have those credences*?

So again, we’re stuck with a piece of math—the relation of O-dominance—without a clear philosophical interpretation. Until we see what O-dominance amounts to, such that it has more normative significance than dominance, we can’t appeal to it as a justification for probabilism or conditionalization.

pected value of an act, given that you won’t take it. But that doesn’t shed any light on the notion of the *value* (simpliciter) of an act at a world where you don’t take it.

6.2 What this shows about epistemic utility functions

There's a further puzzling fact that's brought out by considering the role of O-dominance in accuracy-dominance arguments.

Suppose there are two states and two credal options:

	s_1	s_2
a_1	w_1	w_2
a_2	w_3	w_4

Again, in ordinary dominance arguments, we just care about whether the utility of w_1 is greater than that of w_3 and whether the utility of w_2 is greater than that of w_4 . If so, a_1 dominates a_2 . But in O-dominance arguments, we assess whether the value of a_1 is greater than the value of a_2 at all four outcomes, including, e.g., w_3 .

Why is it possible to assess a_1 at w_3 in epistemic decision theory, given that it wouldn't be possible to do so in practical decision theory? I want to suggest the following answer: each epistemic option generates and is assessed according to a different epistemic utility function over worlds. Scoring rules—rules like the Brier score—aren't sufficient to provide epistemic utility functions over worlds. They have to be paired with credence functions to generate utilities.

As a result, different credence functions, paired with a shared scoring rule, generate different utility assignments over the same worlds. Rather than comparing the epistemic utility of different acts according to a single, shared utility function over worlds, we use different utility functions over worlds for each option. This procedure is, as you might guess, very different from what we do in practical decision theory. One difference is that this procedure involves comparing two different value functions over worlds. And so structurally, comparing different credal acts is no different from interpersonal utility comparison or intertheoretic utility comparison. And this is in itself puzzling: it is part of the decision-theoretic orthodoxy that we cannot make interpersonal or intertheoretic utility comparisons.²⁶

A more interesting difference can be brought out by seeing how this sort of procedure might be used in practical decision theory. In the analogous case, we must choose between two acts that will lead us to have different values, and then assess all of the possible outcomes on the basis of each value system. Suppose I'm deciding whether to take a pill that causes me to value tripping people. I know

²⁶ See e.g. (Robbins, 1932), (Arrow, 1951/1963).

that in the future I'll either trip people frequently or just occasionally, by accident. Suppose also that all that matters to me is how frequently I trip people; that's the sole determinant of value. From the perspective of the values I'll have if I take the pill and come to value tripping people, the values of the outcomes look like this:

	often	rarely
pill	10	4
no pill	10	4

Why isn't the value of (no pill & often) lower? Because from the perspective of the values I would adopt if I took the pill, I still trip people often in that outcome. (I might not enjoy that; but the values I would have if I took the pill only care about my tripping people, not whether I enjoy it.)

Meanwhile, from the perspective of the values I'll have if I continue to disvalue tripping people, the values of the outcomes look like this:

	often	rarely
pill	-10	-4
no pill	-10	-4

If we cannot rationally choose O-dominated outcomes, then we conclude: things look rosier in every outcome from the perspective of the values I'll have if I take the pill, so I should take it. This is obviously not a sensible way of reasoning. All the outcomes should be evaluated relative to a shared utility function that is independent of the acts.

If we stick with this sort of decision-theoretic conception of epistemic utility theory, then there's a natural solution. If we use a utility function that assesses the values of credence functions only at worlds where the relevant agent has those credences, then we can figure out which credences dominate according to a shared utility function.

But I don't think that epistemic utility theorists should take that strategy, since doing so leads to using decision-theoretic rules like Dominance and causal or evidential MaxExAcc that conflict with conditionalization and probabilism. Instead, I think we should clarify the notion of an epistemic utility function. We should

think of epistemic utility functions as ranging over a different kind of formal object. Rather than having the outcomes be propositions (sets of worlds), we can let outcomes be sets of pairs of worlds and credence functions. Then we assess each epistemic act at each world-credence function pair that has that act as its second element.

While I think this strategy is necessary for both O-Dominance arguments and observational MaxExAcc arguments, I want to note that it's quite a hoop to jump through in order to ensure that we can assess the value of credence functions at worlds where the relevant agents don't adopt them. And so if this is what we should do, then we're at least owed a philosophical interpretation of *why* it's crucial to assess credence functions at worlds where no one has them—and in doing so allow e.g. Moore credences not to be accuracy dominated, in spite of the fact that transparency generates a sure gain in accuracy. Otherwise, this is just so much more uninterpreted formalism.

7 Accuracy and evidence

I've argued that the notions of expected accuracy and accuracy dominance that are used in epistemic utility theory don't have a clear philosophical interpretation. And so some important questions about the foundations of epistemic utility theory remain open. By way of conclusion, let me discuss what I think are the main philosophical stakes in answering these questions.

Here are two claims that are pretty plausible:

Evidence You should adopt the belief state that your evidence supports.

Accuracy What you should believe is determined by the aim of accuracy: that is, the aim of having a belief state that's as close to the truth as possible.

These claims seem to go tidily together. It can even seem like Evidence and Accuracy are two different ways of expressing the same thing. After all, the evidence supports believing that *H* only if the evidence suggests that *H* is true. That's what makes it *evidential* support. Some considerations might favor believing what's false, maybe because you derive some benefit from having a false belief. But evidential considerations in favor of a belief are considerations that point you in the direction of the truth, or at least what your evidence suggests is true. If you had perfect, complete evidence, then it would support perfectly accurate beliefs.

But these claims are not the same. There's an intuitive sense in which, in cases like the handstand example, evidential support and promotion of accuracy simply

come apart. Accuracy demands extremal credences in that example. Intuitively, Evidence doesn't.

What does Evidence require in that example? It's not at all clear. Here is a (non-exhaustive) list of some positions one might take with respect to the handstand example:

1. **Impermissivism:** there's a particular credence in H that's rationally required.
2. **Permissivism:** more than one credence in H can be rational, as long as it's arrived at by rational update on permissible priors.
3. **Extreme permissivism:** more than one credence in H is rational, and the choice is unconstrained by your priors.
4. **Gappy credences:** you shouldn't have any credence at all in H .²⁷
5. **Epistemic dilemma:** there are no permissible doxastic states you can adopt.
6. **Belief indeterminacy:** you should adopt an indeterminate doxastic state.
7. **Normative indeterminacy:** it's indeterminate what doxastic state you should have.

For my part, I'm actually somewhat attracted to option 7.²⁸ Your evidence can only support confidence in H to the extent that it supports confidence that you'll believe H . After all, your credence in H is the sole determinant of how likely H is to be true. So if anyone wants to find out whether H , they must investigate how confident you'll be in H . But how do we find out what you'll believe? We've stipulated that you're rational (because we're trying to find out what attitude is rational to take, and we do so by considering what attitudes rational agents take). But wait: if you're rational, you'll be confident of H only on the grounds that the evidence supports H . After all, rationality requires believing H only to the extent that the evidence supports H .

So there is no independent evidential grip we can get on H . Before we can find out what credence is rational in H , we need to find out what credence you have in H (since this determines whether H); but before we can find out what credence you

²⁷ Perhaps because you shouldn't take any attitude toward the proposition that you adopt this or that credence in H . Some ((Levi, 1997), (Spohn, 1977), Briggs (personal communication)) say we shouldn't, or needn't, have credences in our own future acts—at least in the context of deliberation.

²⁸ I defend a similar position with respect to indeterminate evidence in (Carr, manuscriptb).

have in H , we need to know what credence is rational in H (because you are, *ex hypothesi*, rational). There doesn't seem to be an evidential basis for *any* credence here. So it seems plausible to me that the norm of Evidence doesn't give any verdict in this case whatsoever.²⁹

I began by claiming that epistemic utility theory faces a dilemma. We have to choose between traditional decision-theoretic rules (causal/evidential MaxExAcc; Dominance) and observational rules (observational MaxExAcc; O-Dominance). There are costs and benefits associated with each horn.

The benefit of the decision-theoretic rules is that they vindicate the idea that rational belief has the aim of accuracy, and so have an intuitive rationale. Their cost is that we'll have to give up probabilism, conditionalization, and other plausible epistemic norms.

The benefit of observational rules is that we can retain probabilism, conditionalization, etc. Their cost is that we'll have to give up on the idea that epistemic rationality is a matter of pursuing the goal of accuracy—and thereby give up our intuitive explanation for why we should obey these rules.

The second horn is more promising. But it leaves big unopened questions at the foundations of epistemic utility theory.

- Why are worlds where an act isn't taken relevant for the assessment of the act?
- Observational expected accuracy isn't the expected accuracy of an epistemic act or a person's epistemic states; so what is it?
- Why is the accuracy of what OEA assesses epistemically significant, given that the accuracy of a person's epistemic states isn't?

Until these questions are answered, the observational rules stand in need of a philosophical interpretation. Before we can claim that epistemic utility theory justifies

²⁹ Note that even if we add extra evidence—say, that a 80% of people who were in your position and learned $Cr(H) = Ch(H)$ ended up making H true—I think it's rational for that evidence to be screened by $Cr(H) = Ch(H)$, in the sense of Weatherson (manuscript). That is:

1. $Cr(H \mid 80\% \text{ info}) > Cr(H)$
2. $Cr(H \mid 80\% \text{ info} \wedge Cr(H) = Ch(H)) = Cr(H \mid Cr(H) = Ch(H))$

In the light of learning $Cr(H) = Ch(H)$, the information about other people is neutralized. So it's not that you have *no* evidence about H . It's that what evidence you have doesn't seem to make any push toward any credence.

epistemic norms like conditionalization or probabilism, we need to explain and justify the non-consequentialist norms we're treating as fundamental.

Do the observational rules respect the norm of Evidence, instead of Accuracy? If they do, it would provide a special vindication of the research program of epistemic utility theory. Epistemic utility theory would be able to provide something that other forms of epistemic consequentialism so far haven't succeeded in providing: a mathematically precise epistemological theory that codifies the fundamental epistemic norm that we should believe what our evidence supports. The problem is that the arguments that philosophers in epistemic utility theory have provided for observational norms are stated in terms of the goal of accuracy promotion. And so, if the norms associated with OEA and O-dominance make recommendations that align with evidentialism, it's not at all clear *how* or *why* they do. I think these are the questions we should be answering.

Bibliography

- Arntzenius, Frank (2003). "Some Problems for Conditionalization and Reflection." *The Journal of Philosophy*, 100(7): pp. 356–370.
- Arrow, Kenneth (1951/1963). *Social Choice and Individual Welfare*. New York: Wiley.
- Berker, Selim (2013). "Epistemic Teleology and the Separateness of Propositions." *Philosophical Review*, 122: pp. 337–393.
- Berker, Selim (forthcoming). "The Rejection of Epistemic Consequentialism." *Philosophical Issues*.
- Burge, Tyler (1997). "Interlocution, perception, and memory." *Philosophical Studies*, 86: pp. 21–47.
- Caie, Michael (2012). "Belief and Indeterminacy." *Philosophical Review*, 121(1): pp. 1–54.
- Caie, Michael (forthcoming). "Rational Probabilistic Coherence." *Philosophical Review*.
- Carr, Jennifer (manuscripta). "How to Expect When You're Expecting."
- Carr, Jennifer (manuscriptb). "Imprecise Evidence without Imprecise Credences."
- Christensen, David (2000). "Diachronic Coherence Versus Epistemic Impartiality." *Philosophical Review*, 109(3): pp. 349–371.
- Christensen, David (2010). "Rational Reflection." *Philosophical Perspectives*, 24(1): pp. 121–140.
- Elga, Adam (2008). "Lucky to be Rational." (Unpublished).
- Elga, Adam (2010). "Subjective Probabilities Should Be Sharp." *Philosophers' Imprint*, 10(05).

- Elga, Adam (2012). "The Puzzle of the Unmarked Clock and the New Rational Reflection Principle." (Unpublished).
- Friedman, Jane (manuscript). "Inquiry and Agnosticism."
- Gibbard, Allan (1990). *Wise Choices, Apt Feelings: a Theory of Normative Judgment*. Oxford: Clarendon Press.
- Gibbard, Allan (2007). "Rational Credence and the Value of Truth." In *Oxford Studies in Epistemology*, Oxford University Press, vol. 2, pp. 143–164.
- Greaves, Hilary (manuscript). "Epistemic Decision Theory."
- Greaves, Hilary, and David Wallace (2006). "Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility." *Mind*, 115(459): pp. 607–632.
- Hájek, Alan (2012). "Philosophical Heuristics." (Unpublished).
- Hájek, Alan, and Michael Smithson (2012). "Rationality and Indeterminate Probabilities." *Synthese*, 187(1): pp. 33–48.
- Harman, Gilbert (1986). *Change in View: Principles of Reasoning*. MIT Press.
- James, William (1896). "The Sentiment of Rationality." In *The Will to Believe and Other Essays in Popular Philosophy*, Norwood, Mass.: Synthese.
- Joyce, James (1998). "A Nonpragmatic Vindication of Probabilism." *Philosophy of Science*, 65(4): pp. 575–603.
- Joyce, James (2009). "Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief." In *Degrees of Belief*, Synthese, vol. 342, pp. 263–297.
- Joyce, James M. (2010). "A Defense of Imprecise Credences in Inference and Decision Making." *Philosophical Perspectives*, 24(1): pp. 281–323.
- Kelly, Thomas (forthcoming). "How to be an Epistemic Permissivist." In M. Steup, and J. Turri (eds.) *Contemporary Debates in Epistemology*, Blackwell.
- Kyburg, Henry (1983). *Epistemology and Inference*. University of Minnesota Press.
- Leitgeb, Hannes, and Richard Pettigrew (2010a). "An Objective Justification of Bayesianism I: Measuring Inaccuracy." *Philosophy of Science*, 77(2): pp. 201–235.

- Leitgeb, Hannes, and Richard Pettigrew (2010b). "An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy." *Philosophy of Science*, 77(2): pp. 236–272.
- Levi, Isaac (1997). *The Covenant of Reason: Rationality and the Commitments of Thought*. Cambridge University Press.
- Lewis, David (1981). "Causal Decision Theory." *Australasian Journal of Philosophy*, 59: pp. 5–30.
- McGrath, Matthew (2007). "Memory and Epistemic Conservatism." *Synthese*, 157(1): pp. 1–24.
- Meacham, Christopher (manuscript). "Impermissive Bayesianism."
- Meacham, Christopher J. G. (2010). "Unravelling the Tangled Web: Continuity, Internalism, Non-Uniqueness and Self-Locating Beliefs." In *Oxford Studies in Epistemology, Volume 3*, Oxford University Press.
- Moss, Sarah (2012). "Credal Dilemmas." (Unpublished).
- Robbins, Lionel (1932). *An Essay on the Nature and Significance of Economic Science*. London: Macmillan.
- Schoenfield, Miriam (2012). "Chilling Out on Epistemic Rationality: a Defense of Imprecise Credences (and Other Imprecise Doxastic Attitudes)." *Philosophical Studies*, 158: pp. 197–219.
- Schoenfield, Miriam (forthcoming). "Permission to Believe." *Noûs*.
- Seidenfeld, Theodore (2004). "A Contrast Between Two Decision Rules for Use with (Convex) Sets of Probabilities: Γ -Maximin Versus E-Admissibility." *Synthese*, 140: pp. 69–88.
- Skyrms, Bryan (1977). "Resiliency, Propensities, and Causal Necessity." *Journal of Philosophy*, 74: pp. 704–713.
- Spohn, Wolfgang (1977). "Where Luce and Krantz Do Really Generalize Savage's Decision Model." *Erkenntnis*, 11(1): pp. 113–134.
- van Fraassen, Bas (1989). *Laws and Symmetry*. Oxford.
- Weatherson, Brian (2008). "Decision Making with Imprecise Probabilities." (Unpublished).

- Weatherson, Brian (manuscript). "Do Judgments Screen Evidence?"
- White, Roger (2005). "Epistemic Permissiveness." *Philosophical Perspectives*, 19(1): pp. 445–459.
- White, Roger (2009). "Evidential Symmetry and Mushy Credence." In *Oxford Studies in Epistemology*, Oxford University Press.
- Williams, J. Robert G. (2011). "Decision Making Under Indeterminacy." Manuscript.
- Williamson, Timothy (2000). *Knowledge and its Limits*. Oxford University Press.
- Williamson, Timothy (2007). "Improbable Knowing." (Unpublished notes).
- Yalcin, Seth (forthcoming). "Bayesian Expressivism." *Proceedings of the Aristotelian Society*.